HELLENIC REPUBLIC
National and Kapodistrian University of Athens
EST. 1837

DEVELOPMENT OF A SHEAR WAVE SPLITTING PARAMETERS DETERMINATION SOFTWARE SUITE FOR UPPER CRUST SEISMIC ANISOTROPY STUDIES: APPLICATION IN TECTONIC AND VOLCANIC REGIMES

by
Ioannis Spingos

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Seismology

Examining Committee:
Dr Kaviris George (Supervisor, Assistant Professor)
Dr Papadimitriou Panayotis (Professor)
Dr Voulgaris Nicholas (Professor)

Athens
2019

Preface

This work is the result of years of efforts in studying shear – wave splitting phenomena and developing my programming skills. It was a personal passion project that started near the end of my undergraduate thesis back in 2015 and morphed into a proper scientific project during my postgraduate studies. After spending countless hours analyzing waveforms using PostScript files and Shell Scripts on a Linux terminal, my intention was to create a modern program that would incorporate contemporary software tools, severely reduce the work hours required and smooth out the process for the uninitiated. The program presented in this MSc Thesis has taken shape and can already be used in research. Nevertheless, progress cannot be halted. New features and updates will continue to be added to enhance the user experience, for (hopefully) years to come.

Every project like this, albeit written and presented by one person, includes several people that influenced it, either in a major or a minor way. Thus, I would like to thank, first and foremost, my supervisor Assistant Professor George Kaviris for the endless hours it took to discuss (more like debate) this project and his illogical faith in my embryonic programming skills at the beginning. I would also like to express my gratitude to the other two members of the Examining Committee, Professor Panayotis Papadimitriou, for providing the necessary stimulus to work on programming, while always pushing me to improve and having a different outlook on things, and Professor Nicholas Voulgaris, for the long years of cooperation in seismic anisotropy. It would be impossible to not acknowledge the aid of fellow MSc student Christos Millas in testing the software and providing valuable insight, through arguments upon arguments, on shear – wave splitting. I am also grateful to Associate Professor Vicky Kouskouna and Assistant Professor Ioannis Kassaras for the priceless knowledge on seismology that they communicated to me, since my early undergraduate years. I would also like to thank Dr Vasileios Kapetanidis, who always set the bar for me in both software development and seismic anisotropy topics, as well as Dr Andreas Karakonstantis and PhD candidate George Bozionelos for the massive undertaking of picking phases and locating earthquakes in the Gulf of Corinth and Santorini island, used in the current thesis. I want to express my thanks to PhD Candidate Phaedra Kravvariti for the manual analysis of the Santorini Volcanic Complex data for shear – wave splitting. I'm also grateful to Dr George Sakkas, PhD candidate Nikos Sakellariou and fellow MSc student Ioannis Fountoulakis for our everyday interaction and support through my years in the Seismological Laboratory. I would like to thank two members of the University of Kansas, Professor George Tsoflias and PhD candidate Alex Keith Nolte for providing data from the Wellington Oil Field in Kansas and all the ideas that popped up during our discussions about shear – wave splitting. I am greatly thankful to fellow undergraduate student and physicist Mr. Christoforos Vasileiou for the ceaseless support and motivation he provided all these years. Lastly, but definitely not least, I'm very grateful to my sister Dr Foteini Spingou and my parents, Katerina and Andreas, for their ever – going support, both financially and psychologically, through my studies.

Contents

Περίληψη

Το φαινόμενο της σχάσης (κατά το οποίο ένα εγκάρσιο κύμα διακρίνεται σε δύο συνιστώσες με κάθετες μεταξύ τους πολώσεις και διαφορετικές ταχύτητες διάδοσης) έχει αποδοθεί στην ανισοτροπία του μέσου διάδοσης. Η παρατήρησή του εξαρτάται από τον προσδιορισμό δύο κύριων παραμέτρων: τη διεύθυνση πόλωσης της ταχείας συνιστώσας (φ) και της χρονικής καθυστέρησης μεταξύ των χρόνων διαδρομής (td). Η παρούσα Μεταπτυχιακή Διατριβή Ειδίκευσης αφορά την ανάπτυξη ενός πακέτου λογισμικού (σε γλώσσα Python) για τον προσδιορισμό των παραμέτρων σχάσης. Αν και υπάρχουν διάφορες εφαρμογές διαθέσιμες, σε καμία μέχρι σήμερα δεν προσφέρεται η ολοκληρωμένη χρήση χειρακτικών, ημι- και πλήρως αυτόματων μεθόδων. Το λογισμικό το οποίο αναπτύχθηκε (Pytheas) σκοπεύει να εναρμονίσει τη λειτουργία των τεχνικών αυτών. Η φιλικότητα προς τον χρήστη είναι ένα από τα βασικά χαρακτηριστικά του προγράμματος. Το πρόγραμμα εμπεριέχει τρεις μεθόδους προσδιορισμού: (α) εποπτική μέθοδος Διαγραμμάτων Σωματιδιακής Κίνησης, (β) μέθοδος Ιδιοτιμών και (γ) μέθοδος Περιστροφής – Ετεροσυσχέτισης. Ενώ η (α) είναι χειρακτική μέθοδος και απαιτεί την πλήρη ενασχόληση του χρήστη, οι (β) και (γ) είναι ημι – αυτόματες και χρειάζονται τον χειρακτικό προσδιορισμό του παραθύρου στο οποίο θα γίνει η μέτρηση. Η επιλογή αυτή αυτοματοποιείται πλήρως με τη μέθοδο Ανάλυσης Συστάδων, σε συνδυασμό με τη (β) ή και τη (γ). Το λογισμικό Pytheas συνδέει τις τεχνικές αυτές με ένα φιλικό και εύληπτο ενσωματωμένο γραφικό περιβάλλον χρήστη. Για την ορθή αξιολόγηση των μετρήσεων, αναπτύχθηκε ένα σύστημα βαθμολόγησης στηριζόμενο στο περιεχόμενο του θορύβου στο σήμα και στα σφάλματα των δύο παραμέτρων σχάσης.

Για τον έλεγχο της λειτουργίας του Pytheas χρησιμοποιήθηκαν τοπικά δεδομένα από τρεις περιοχές: (i) Δυτικός Κορινθιακός Κόλπος (τεκτονικό περιβάλλον), (ii) Ηφαιστειακό Σύμπλεγμα Σαντορίνης (ηφαιστειακό περιβάλλον) και (iii) Πετρελαιοφόρο πεδίο Wellington, Kansas, USA (επαγόμενη σεισμικότητα). Η Ανάλυση Συστάδων χρησιμοποιήθηκε για την επιλογή του ιδανικού παραθύρου ανάλυσης και η μέθοδος Ιδιοτιμών εφαρμόστηκε σε αυτό, για τον προσδιορισμό των παραμέτρων σχάσης σε πάνω από 2,400 ζεύγη σεισμών – σταθμών, με γωνίες ανάδυσης μικρότερες των 35°. Προσδιορίστηκαν συνολικά 516 ζεύγη παραμέτρων σχάσης με αποδεκτούς βαθμούς ποιότητας. Τα αποτελέσματα, συγκρινόμενα με αντίστοιχα παλαιών μελετών με τη χειρακτική μέθοδο, έδειξαν ότι η αυτόματη μέθοδος έχει σημαντική αξιοπιστία. Ωστόσο, η πιο έντονη διασπορά των τιμών των φ και td θα πρέπει να διερευνηθεί, πιθανώς με τη βελτιστοποίηση της αυτόματης διαδικασίας βαθμολόγησης.

Επιπλέον, εξετάσθηκε η μέθοδος της Σεισμικής Τομογραφίας Σχάσης Εγκαρσίων Κυμάτων, με τη χρήση μετρήσεων παλαιών ερευνών (λόγω του μεγαλύτερου πλήθους τους), στον Δυτικό Κορινθιακό Κόλπο και το Ηφαιστειακό Σύμπλεγμα Σαντορίνης. Τα αποτελέσματα έδειξαν ότι απαιτείται περαιτέρω πληροφορία για την ερμηνεία της τομογραφίας, καθώς οι προσδιορισθείσες ανισοτροπικές δομές του ανώτερου φλοιού δεν συνάδουν με γνωστά στατικά γεωλογικά και τεκτονικά φαινόμενα. Η διερεύνηση της κατανομής της τάσης στην περιοχή (μέσω μετρήσεων πίεσης σε γεωτρήσεις ή αντιστροφής μηχανισμών γένεσης) θα μπορούσε να χρησιμοποιηθεί για τον σκοπό αυτόν, επιτρέποντας τη συσχέτιση της τομογραφίας με δυναμικά φαινόμενα.

Abstract

The splitting phenomenon (during which a shear – wave is divided in two orthogonally polarized components travelling with different velocities) is attributed to the anisotropy of the propagation medium. Its observation is relying on identifying two main parameters: the polarization direction of the fast shear – wave (φ) and the time – delay between the travel – times of the two components (t_d). The current MSc thesis aims to develop a software suite (in the Python programming language) for determining the splitting parameters from local events, in the upper crust. Even though there are various programs available, they do not offer the integrated use of manual, semi- and fully automatic methods. The developed software (Pytheas) aims to harmonize the implementation of these techniques. User - friendliness is a basic feature of the program. The application incorporates three methods: (a) visual inspection of particle motion diagrams, (b) eigenvalue method and (c) rotation – correlation method While (a) is a manual technique and requires the complete involvement of the user, (b) and (c) are semi – automatic and depend upon the manual determination of an analysis window. This selection is fully automated by using Cluster Analysis, in combination with either (b) or (c). The Pytheas software connects all these techniques through a friendly and intuitive graphical user interface. For the evaluation of the measurements, a grading system was developed, based on the noise in the signal and the errors of the two splitting parameters.

For the evaluation of Pytheas, local data from three different areas were used: (i) the Western Gulf of Corinth (tectonic regime) in Greece, (ii) the Santorini Volcanic Complex (volcanic regime) in Greece and (iii) the Wellington Oil Field in Kansas, USA (induced seismicity). Cluster analysis was utilized in selecting the optimal analysis windows for the eigenvalue method, in over 2,400 event – station pairs, with a maximum angle of incidence equal to 35°. A total of 516 measurements with acceptable grades were determined. Results, compared with corresponding ones from previous studies that used the manual method, showed that the automated scheme is significantly reliable. However, the more intense scatter in both φ and t_d must be investigated, possibly by refining the automatic grading algorithm.

In addition, the Shear – Wave Splitting Tomography method was examined, by using manual measurements of past studies (due to their increased number of observations) in the Western Gulf of Corinth and the Santorini Volcanic Complex. Results showed that, to adequately interpret the tomography, additional information is required, given that the determined anisotropic structures do not agree with static geologic and tectonic features of the upper crust. The investigation of the stress distribution (by pressure measurements in wells or the inversion of focal – mechanisms) could be used towards this goal, to relate the tomography to dynamic phenomena.

1. Introduction

In Physics, anisotropy constitutes the directional dependence of a physical quantity. From computer graphics to neuroscience, anisotropy is a significant property of any material. Its characteristics are exploited, whether to examine the aging process of the brain (Salat, 2014) or to investigate the further reaches of the cosmos (Smoot et al., 1977). In Seismology and Geophysics, anisotropy has been a known property of the propagation medium, since their early stages. In this Chapter, the characteristics of Seismic Anisotropy and its applications in modern literature will be examined.

1.1. Seismic Anisotropy in Shear – Waves

The concept of Seismic Anisotropy refers to the variation of the velocity of seismic waves in different directions, caused by inhomogeneities in the propagation medium. Even though anisotropy is a property of the medium and, hence, affects both body and surface waves, the most prevalent associated phenomenon is Shear – Wave Splitting (SWS).

SWS refers to the observation that when a shear – wave enters an anisotropic medium, two components of polarization can be distinguished. The one travelling with the higher velocity is generally polarized according to the main feature of anisotropy (S_{fast}). The second showcases a polarization direction perpendicular to the above (S_{slow}). SWS studies are based on the determination of two parameters: the polarization direction of S_{fast} (φ) and the time – delay between the arrivals of the two split shear – waves at the station (t_d). Crampin et al. (1984) presented a concise review of the – then – current knowledge of seismic anisotropy and its effects. Following, a summary of their revision is presented.

For the fourth – order elastic tensor C_{jkmn} and the displacement u_j in direction j, the equation of wave propagation is:

$$\frac{\rho \partial^2 u_j}{\partial t^2} = C_{jkmn} u_{m,nk} \tag{1.1}$$

where ρ the density, $u_{m,nk} = \frac{\partial^2 u_m}{\partial x_n \partial x_k}$ and the directional indices (j, k, m, n) of the plane x belong in the $[1,2,3]$ range (equivalent to the x, y and z axes). Let a plane x_p of mirror symmetry in the direction p. Reflections on x_p will not affect the elastic constants, due to the symmetry assumption. Hence, for the reflected elastic tensor C'_{jkmn} the following conditions exist:

$$\begin{aligned} C'_{jkmp} &= -C_{jkmp} \\ C'_{jkpp} &= C_{jkpp} \\ C'_{jpmp} &= C_{jmpm} \\ C'_{jppp} &= -C_{jppp} \\ C'_{pppp} &= C_{pppp} \end{aligned} \tag{1.2}$$

Thus, the elastic constants remain the same when $C_{jkmn} = 0$ and one or three of the directions are equal to p. Considering Eq. (1.2), we can derive from Eq. (1.1):

$$\begin{aligned} \rho c^2 a_1 &= C_{1111}a_1 + C_{1121}a_2 + C_{1131}a_3 \\ \rho c^2 a_2 &= C_{2111}a_1 + C_{2121}a_2 + C_{2131}a_3 \\ \rho c^2 a_3 &= C_{3111}a_1 + C_{3121}a_2 + C_{3131}a_3 \end{aligned} \tag{1.3}$$

where c is the phase velocity and $a_j = \frac{u_j}{e^{i\omega(t-\frac{x_1}{c})}}$. The common multiplier $(-i\omega)^2 e^{i\omega(t-\frac{x_1}{c})}$ is omitted. The above equations can be solved if considered as a linear eigenvalue problem. Eq. (1.3) can be written as:

$$(\boldsymbol{T} - pc^2\boldsymbol{I})\boldsymbol{a} = 0 \tag{1.4}$$

where $\boldsymbol{I}$ is the 3 x 3 identity matrix, $\boldsymbol{a}$ is comprised by elements a_j and $\boldsymbol{T}$ is a 3 x 3 matrix containing the elastic constants of Eq. (1.3). The problem presented in Eq. (1.4) has three real positive roots for ρc^2, which correspond to three different body – wave types; the quasi P – wave (qP) and two quasi S – waves (qS_1 and qS_2, or, as mentioned above, S_{fast} and S_{slow}). All components are polarized in mutually orthogonal directions. Even though all three components should exist when propagation occurs in an anisotropic solid, the qP energy is negligible. The described behavior refers to the anisotropic part of the ray path. When the, now split, shear – waves exit the anisotropic medium and traverse an isotropic one anew, the difference in velocities is not retained and both qS_1 and qS_2 travel with the same velocity. As a result, the modification of the original pulse due to anisotropy is still present.

Concerning the cause, or origin, of anisotropy in the Earth's interior, a wide variety of models have been proposed. The unequivocally most dominant model on SWS in the crust describes the existence of fluid – filled microcracks, aligned according to the regional stress regime. Cracks are ubiquitous in the brittle rocks of the Earth's interior. There is a wide variety of crack types (Table 1.1), but they all originate in inherent regions of rock weakness. Such areas, susceptible to failure, are flaws, pores and grain contacts. Crack initiation and propagation is directly linked to the stresses exerted on a rock volume and the intrinsic characteristics of the rock, such as porosity and rock fabric (Dresen and Guéguen, 2004). In seismic zones and active faults, microcracks have been linked to fluid migration mechanisms, either related to pre – seismic fault loading or post – seismic fluid discharge (Sibson, 1994).

Table 1.1 Top: Basic crack categories and their origin (Dresen and Guéguen, 2004). Bottom: Crack propagation (or fracture) modes (Sun and Jin, 2012).

Type	Origin	Description
Intergranular	Stress induced, thermal expansion, elastic differences among mineral phases	They develop on grain boundaries and lead to grain – grain or grain – cement separation.
Intragranular	Grain boundaries, pores or inclusions inside grains	Can extent to multiple grains (*transgranular cracks*). Usually lead to arrays of parallel *cleavage cracks.*
Mode	**Process**	**Stress Behavior**
I	Opening	Tensile stress normal to the crack plane
II	Sliding	Shear stress parallel to the crack plane, but perpendicular to the crack front
III	Tearing	Shear stress parallel to both the crack plane and front

Proposed by Crampin (1978) and Crampin et al. (1984), the Extensive – Dilatancy Anisotropy (EDA) model suggests that the properties of tensile cracks in rock solids are greatly affected by the maximum horizontal stress component ($\sigma_{H\,max}$). Over long time periods, the stresses accumulated in a rock volume (associated with earthquake preparatory processes, but not constrained in the immediate focal zone) cause the corrosion of existing fractures, in tandem with water, at slow rates. Furthermore, microcracks undergo a process of elastic bowing, when pore pressure is high enough to act as an internal stress agent. The differential opening (and, consequently, closing) of microcracks leads to a structure that is pervaded by these vertical inhomogeneities and causes SWS to propagating shear – waves. The polarization direction of the S_{fast} component is aligned per the $\sigma_{H\,max}$. According to Crampin et al. (1984), EDA is mainly controlled by four factors: (i) the affinity of the randomly distributed microcracks for opening and closing, (ii) the water existence probability in cracks throughout the crust, (iii) the corrosive potency of stress that leads to crack growth and weakening and (iv) the effective anisotropy of dilatant rocks. As described above, in the case of a single anisotropic layer, the polarization of the S_{fast} will be parallel to the crack plane and two split shear – waves will be observed. However, if the ray passes through a second anisotropic layer with a different microcrack orientation, the polarization will change accordingly and splitting will occur again, leading to four split waves (a pair for each of the original S_{fast} and S_{slow}). This mechanism obfuscates any information about the deeper layer (Yardley and Crampin, 1991). EDA has been used to interpret the anisotropic structure in a variety of areas worldwide, such as the

seismically inactive region of northern Scandinavia (Brooks et al., 1987) and the active rift of the Gulf of Corinth in Greece (Bouin et al., 1996; Papadimitriou et al., 1999).

The main drawback of the EDA model was the inability to interpret variations in the measured anisotropy parameters. If the S_{fast} polarization is aligned according to $\sigma_{H\,max}$, where does the variability of observed polarization directions come from, barring any technical issues (e.g. data noise or free surface interactions)? Microcracks are known to evolve in time and constitute a continuously changing phenomenon (Dresen and Guéguen, 2004). Such observations implied a dynamic and responsive anisotropic behavior of shear – waves and EDA matured into the Anisotropic Poro – Elasticity (APE) model (Crampin and Zatsepin, 1997; Zatsepin and Crampin, 1997). In APE, pore fluid pressure (p_f) is a major factor in the SWS phenomenon. The difference, called effective stress, between p_f and the minimum horizontal compression ($\sigma_{H\,min}$) controls the opening (negative effective stress) and closing (positive effective stress) of the microcracks. Given that $\sigma_{H\,min}$ is generally stable for an area, the effective stress is essentially affected exclusively by p_f. Thus, SWS is sensitive to pore fluids and their features (temperature and composition). In more detail, Zatsepin and Crampin (1997) developed a model that considers SWS from the perspective of microcrack evolution. The deformation of any drained crack volume V can be expressed as:

$$\frac{V}{V_0} = 1 + \frac{\sigma_n}{\sigma^c} \tag{1.5}$$

where V_0 the initial volume, σ_n the normal applied stress and $\sigma^c = \frac{1}{c_{cr}}$. In turn, c_{cr} is the crack compressibility defined as:

$$c_{cr} = \frac{1}{V_0}\frac{\partial V}{\partial \sigma_n} \tag{1.6}$$

From Eq. (1.5) it is evident that the closing of the crack begins when $|\sigma_n| > |\sigma^c|$. In essence, σ^c constitutes the stress boundary between open and closed microcracks. This is true for those oriented according to $\sigma_{H\,max}$. For a fluid – saturated microcrack, a similar expression to Eq. (1.5) is:

$$\frac{V}{V_0} = 1 + \frac{\sigma_n + p_f}{\sigma^c} \tag{1.7}$$

Considering that σ_n can be written as a function of the angles ψ (between the crack normal and the vertical axis x_3) and θ (between the horizontal projection of the crack normal and horizontal axis x_1), for a triaxial system, as:

$$\sigma_n = \sigma_1 \sin^2\psi \cos^2\theta + \sigma_2 \sin^2\psi \sin^2\theta + \sigma_3 \cos^2\psi \tag{1.8}$$

it is evident that the crack orientation is a crucial factor regarding both the evolution of the microcracks and the σ^c. The perturbation of the aspect ratio (g) of any microcrack is given by:

$$\frac{g(\sigma_n)}{g_0} = 1 + \left(\frac{2(1-v)}{\pi g_0 \mu}\right)\sigma_n \tag{1.9}$$

where μ the shear modulus, ν the Poisson ratio and g_0 the initial aspect ratio. Given that $\frac{V}{V_0} \cong \frac{g}{g_0}$, the orientation of the microcracks and pore fluid pressure also affect the crack geometry. If Eq. (1.7) is defined as the function $\gamma(\sigma_{ij}, p_f, \psi, \theta)$, with σ_{ij} equivalent to σ_n, we can infer that γ is responsible for the opening and closing of the microcracks. Zatsepin and Crampin (1997) proposed a relationship to estimate the characteristic time of fluid diffusion due to stress (t^c). Expressed as a function of σ^c, this quantity is described by:

$$t^c \propto \frac{5\left(\frac{L}{a}\right)^3 \left(\frac{\sigma_c}{\mu}\right) a^2 \eta}{k \Delta\sigma} \tag{1.10}$$

where L the average crack spacing, α the crack radius, η the pore fluid viscosity, k the permeability and $\Delta\sigma$ the deviatoric component of the differential stress field. Figure 1.1 showcases some examples of expected diffusion times. Even though every rock mass has different hydraulic properties, Eq. (1.10) provides a robust reference point for expected temporal variations of the anisotropic properties of shear – waves. If a fluid requires a couple of hours to migrate from an intergranular crack to another, it is not possible to observe SWS changes in a narrower timeframe.

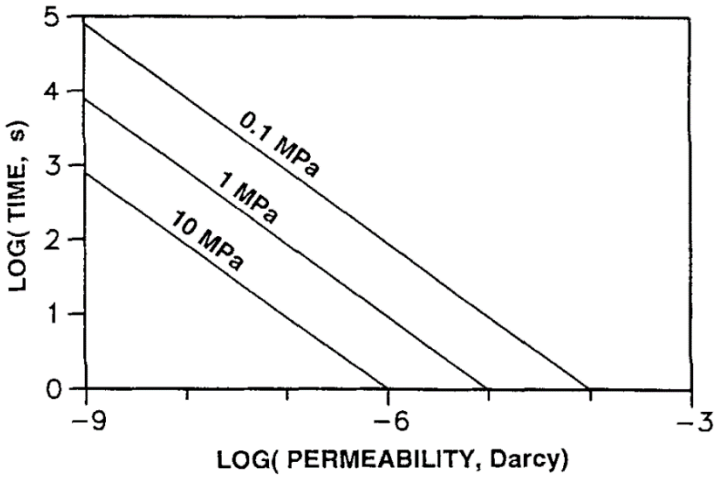

Figure 1.1. Characteristic fluid diffusion times obtained from Eq. (1.10) for various $\boldsymbol{\Delta\sigma}$ and $\boldsymbol{k}$ values (Zatsepin and Crampin, 1997). For the lowest permeabilities, the diffusion time ranges from 10^3 to 10^5 s, i.e. between 0.3 and 28 h.

Nevertheless, shifts in stress may take less time than t^c. In that case, the crack volume is responsible for the SWS variations. Finally, the relationship between SWS and stress / pore dynamics is given by:

$$C_{ijkl} = C^s_{ijkl} + \varepsilon_0 \iint_{\gamma>0} R^{pqrs}_{ijkl}(-\psi, -\theta) C^1_{pqrs}(\gamma, r) d\Omega \tag{1.11}$$

where C_{ijkl} the elastic tensor, C^s_{ijkl} the elastic tensor of the solid matrix, ε_0 the effective crack density (< 0.1), R the rotation matrix from $ijkl$ to $pqrs$, C^1_{pqrs} the first – order correction for elastic constants of parallel cracks and r the normalized pore fluid compressibility. It is evident from Eq. (1.11) that stress variations in the rock mass are indeed reflected in SWS. To summarize, the formulations that constitute the APE model showcase the impact of the state of stress through: (i) the changes in the microcracks' distribution, volume and, especially, aspect

ratio and (ii) the effect of fluids, either because of their intrinsic properties (pressure, temperature, composition) or due to diffusion processes. Another critical consequence of the above is the effect of high pore fluid pressure. When that occurs, the polarization direction of the S_{fast} can shift from parallel to perpendicular to σ_{Hmax} (essentially, the would – be S_{slow} component attains higher velocity). This phenomenon has been described as "90° flip". The APE model paved the way for the plausible monitoring of stress cycles through processing of seismological data (discussed in further detail in Chapter 1.2), rendering SWS changes as potential earthquake or volcanic precursors. Crampin and Zatsepin (1997) discussed this possibility and concluded, based on observations and theoretical modelling, that APE satisfyingly explains the anisotropic behavior of shear – waves in the upper crust (up to a depth of 15 km) and can be used to explain fluid – rock interactions in that context. APE has been applied extensively to interpret SWS in cracked media since the late 90s (e.g. Crampin et al., 1999; Bianco et al., 2006; Gao and Crampin, 2006; Kaviris et al., 2015, 2017, 2018a, 2018b).

Seismic anisotropy in the crust has also been attributed to sources other than fluid – saturated microcracks. Sedimentary rocks present various strongly anisotropic features (e.g. bedding and fine layering). In cases where the mineralogy of the specimen is dominated by phyllosilicates (e.g. shales), the Lattice – Preferred Orientation (LPO) of those minerals greatly enhances SWS. Quartz – rich rocks do not present such features. This is because the phylosilicates constitute a stronger LPO. Inferred SWS values in shales showcased almost ten times stronger anisotropy compared to sandstones (Valcke et al., 2006). LPO is commonly used to explain the splitting of teleseismic phases (e.g. SKS). In that case, the anisotropy is caused by the alignment of the olivine crystals in the upper mantle (e.g. Vinnik et al., 1989). Another layer – related source of seismic anisotropy is the Periodic Thin Layering (PTL). Stratigraphies that are comprised of multiple isotropic layers can lead to an effect similar to LPO and cause SWS, even though there are several restrictions due to the split wave's wavelength (Helbig, 1984).

1.2. Recent Developments in Geophysics and Seismology

A direct consequence of the APE model is the stress monitoring through SWS observations, especially of t_d. Crampin et al. (1999) achieved in 1998 in Iceland an *a priori* "stress – forecast", meaning the determination of a possible range of magnitudes and origin times of an impending earthquake. They used variations of normalized according to the hypocentral distance time – delays (t_n) to infer the accumulation and subsequent release of stress, leading to a M = 5.0 event. An initial warning was sent two weeks before the occurrence to the Icelandic authorities, with the stress – forecast finalized three days prior. This experience provided hope for a future robust earthquake prediction tool. However, even though such "forecasts" of seismic (and volcanic) events have been repeated, they all have been in hindsight (Bianco et al., 2006; Gao and Crampin, 2006; Crampin et al., 2015). It has even been claimed that SWS could be sensitive to changes in a tectonic plate scale for very large earthquakes, as in the case of the 2004 Sumatra M_w = 9.2 event (Crampin and Gao, 2012). Regarding the mechanism of a stress – forecast, coming from APE, the increase of t_n, a consequence of microcracks opening, should correspond to stress loading in the vicinity of the seismological station. In a shorter time-period before the event, t_n should start dropping. This reduction is equivalent to a stress drop caused by the coalescence of microcracks to the future fault plane. When t_n reaches a critical threshold,

failure occurs and the earthquake takes place (Crampin, 2011; Crampin et al., 2013). The stress accumulation and release periods range from hours to years (Figure 1.2), depending on the magnitude of the forthcoming earthquake. For a M = 3.0 event, the increase is expected to last 17 days and the decrease about 3 hours. Similarly, the respective values for a M = 7.0 earthquake are 3 years and 1 year (Crampin et al., 2008).

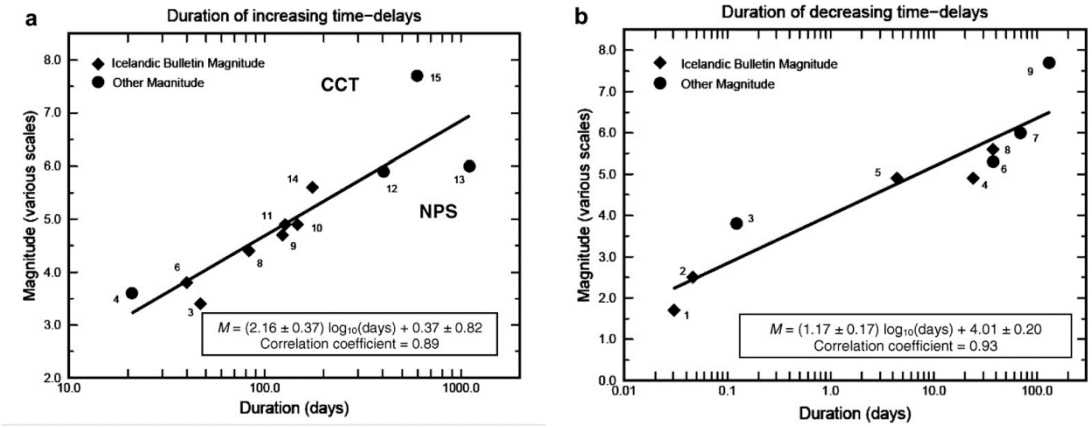

Figure 1.2. Time periods required for t_n increase (a) and decrease (b) as a function of magnitude, after Crampin et al. (2008). Point labels refer to earthquakes cited in the article. CCT (No 15 in (a) and 9 in (b)) is the 1999 M_w = 7.7 Chi – Chi earthquake in Taiwan and NPS (No 13 in (a) and 7 in (b)) is the 1988 M_s = 6.0 North Palm Springs, California (USA) event. The linear relationships between duration and magnitude present strong correlation with observations.

Nevertheless, there are several conundrums regarding this approach. A universal issue in SWS studies is the wide scatter of time – delays that does not permit the extrapolation of robust increase and decrease trends. Crampin et al. (2004) argued that this behavior can be attributed to high pore fluid pressures causing 90° flips (even if those are observed in a very small number of measurements) and suggested controlled source methods, where no such phenomena are observed. Nevertheless, Kaviris et al. (2017, 2018a, 2018b) did not observe 90° flips (limited or not) in most of the stations studied, while the scatter in time – delays was high. In situations of seismic swarms and intense seismicity, it could be argued that variations of parameters corresponding to the closely (temporally) occurring events overlap and cause confusion. For example, if two moderate earthquakes take place in the span of a couple of days, there is not enough time to distinguish between the stress cycles of the two. The increase of the second event will overlap with the decrease period of the first, leading to both high and low time – delays at the same time. Another potential cause for this could be the analysis method. It has been reported that automatic methods present problems (Crampin and Gao, 2006a). However, even manual methods are not errorproof. Time – delay determination is susceptible to analyst's bias and, no matter how strict the criteria, uncertainties arise from phenomena like the interaction of shear – waves with the free surface (Evans, 1984) and the interference of secondary phases (e.g. SP) that could lead to erroneous identification of direct shear – waves. These are challenges that will continue to puzzle SWS researchers, but if they are resolved, splitting could be able to be used as a robust forecasting tool.

The study of the seismic anisotropy of shear – waves is also valuable in the oil industry, both in exploration and production. Data obtained from controlled source experiments need to be corrected for anisotropy. These corrections lead to the improvement of the resolution of seismic tomography, revealing detailed structures. In addition, the porosity of the propagation

medium is investigated with data of the split reflected phases PS_1 and PS_2 (Bale et al., 2009). During production and accompanying operations, numerous earthquakes, usually of small magnitude, are generated. This induced seismicity can be used in SWS studies. Several researchers utilize recordings of these events to characterize fractures in reservoirs through seismic anisotropy (Al-Harrasi et al., 2011). In Kansas, in the United States of America (USA), wastewater injections have possibly caused an immense increase in seismicity and the occurrence of felt earthquakes. SWS observations in the area have shown correlation with pressure measurements (Nolte et al., 2017). Finally, the strength of anisotropy has been linked to increased production in hydraulic fracturing sites, but not in classic wells (Sandanayake and Bale, 2011).

1.3. The Aim and Contribution of the MSc Thesis

There is a plethora of publicly available software for SWS analysis, such as SplitLab (Wüstefeld et al., 2008), mFast (Savage et al., 2010) and SplitRacer (Reiss and Rümpker, 2017), to name but a few. Even though the available software catalogue covers (more than) adequately the most commonly used semi – automatic and automatic techniques, there are still several ways to improve. The "PYThon sHear wavE Analysis Suite" (Pytheas) software, developed in the framework of the current MSc Thesis, is built on the following axes:

i. Provide an open – source code in Python, an ever – growing popular programming language in science.
ii. Increase accessibility by not requiring proprietary software (e.g. MATLAB) or involve an arduous installation and use process (e.g. most programs written in Fortran and may use Shell Scripts to operate).
iii. Be operating system independent (again, no Shell Scripts that prohibit application in Windows systems or Batch Scripts that exclude Unix users).
iv. Provide a modern, user – friendly Graphical User Interface (GUI) to not intimidate potential users that are not familiar with command prompts and terminals.
v. Offer an integrated solution for the seamless application of manual, semi – automatic and fully automatic methods.
vi. Take advantage of free Python packages and increase functionality by including several secondary Quality of Life (QoL) features.
vii. Be modular, to facilitate updates (whether by the original author or other users).

In addition to the development of Pytheas, the present MSc Thesis tackles several SWS related topics. By applying a fully automatic method to analyze waveforms, this Thesis aims to explore the boundaries of the technique in the Western Gulf of Corinth (WGoC) and the Santorini Volcanic Complex (SVC), in Greece.

Finally, previously manually analyzed measurements will be used to apply the newly developed method of Shear – Wave Splitting Tomography (SWST), to explore the lateral boundaries of seismic anisotropy in the WGoC and SVC.

2. Shear – Wave Splitting Analysis Methods

Since the early 80s, a variety of SWS analysis methods have been proposed and implemented. Whether fully (manual methods), partially (semi – automatic) involving an analyst or not requiring any human interaction at all (automatic), these techniques have presented several disadvantages on their own. Following, a comprehensive description of the analysis techniques implemented in Pytheas is presented. Three SWS analysis techniques are included: (a) the visual inspection of particle motion diagrams, which requires a dedicated analyst (Chapter 2.1), (b) the rotation – correlation (Chapter 2.2) and (c) the eigenvalue (Chapter 2.3) methods. The latter two depend upon the manual selection of the time window where the analysis will take place and are, as a result, characterized as semi – automatic methods. This time window selection can be automated with the employment of cluster analysis algorithms, as shown in Chapter 2.4. Finally, the relatively new technique of SWST will be described in Chapter 2.5.

2.1. Particle Motion Diagrams

Diagrams that exhibit the ground particle motion can be constructed from any two recordings of different components. Regarding shear – waves, the polarization is identified in the two horizontal components, i.e. the North – South (NS) and East – West (EW). As a result, to examine the SWS phenomenon, the diagrams capture the motion in the horizontal North – East plane (NE). Thus, each vector direction (θ_{NE}) can be acquired from:

$$\theta_{NE} = \tan^{-1}\frac{NS}{SW} \tag{2.1}$$

where NS and SW the vector magnitudes in the respective direction (i.e. the amplitude of ground motion in each component). To obtain the magnitude of each vector in the NE plane the following relationship is used:

$$NE = \sqrt{NS^2 + EW^2} \tag{2.2}$$

Any particle motion is the synthesis of the NS and EW vectors in time, as acquired from Eq. 2.1 and Eq. 2.2. There are two categories of diagrams used in SWS analysis. The hodograms (Figure 2.1) present continuous time lapses of the particle motion.

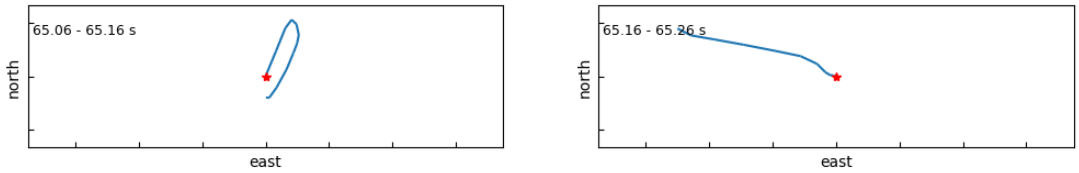

Figure 2.1. Example of two hodograms. The continuous lines are composed of different vectors of particle motion in the NE plane, in the timeframe denoted in the upper left corner of each plot. Even though a sense of time exists, the analyst cannot confidently determine the exact point at which each separate motion begins or ends.

Polarigrams (Bernard and Zollo, 1989) showcase each vector in respect to time (Figure 2.2), in a manner similar to waveforms. Both plots are important in SWS analysis. While polarigrams are, in general, more convenient for the analyst to use, hodograms are utilized when the φ value is close to N90°E and the exact polarization direction is not clear.

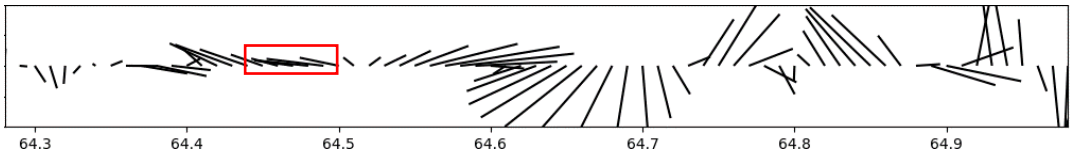

Figure 2.2. Example of a polarigram. Each vector corresponds to the composition of the NS (y axis) and EW (x axis) ground amplitudes. A special case is highlighted (red rectangle) where, if the analyst wishes to determine the polarization direction, the polarigram is not a suitable method and the hodogram is used instead. The time axis is in s.

Particle motion diagrams are visually inspected by an analyst to measure the desired SWS parameters (φ and t_d). The processing scheme used in the current thesis has been previously applied in several case studies in Greece (Papadimitriou et al., 1999; Kaviris et al., 2008, 2015, 2017, 2018a, 2018b). It is noted that Pytheas is the only publicly available software that offers a work environment for this method, as of today. The following example showcases each step of the analysis, as performed in the above literature. Waveforms were recorded in station KALE in the WGoC from an event that occurred on 2013-06-04T22:56:24.000 GMT with a magnitude of M_L = 1.4. The epicentral distance is 9.4 km, while the angle of incidence is equal to 37°. To avoid interactions with the free surface that can lead to misidentification of the shear – wave arrival (e.g. arrival of phases like sP), Evans (1984) and Booth and Crampin (1985) proposed the concept of the shear – wave window, a critical angle of incidence (θ_{crit}) defined by:

$$\theta_{crit} = \sin^{-1}\frac{V_s}{V_p} \tag{2.3}$$

where V_p and V_s the velocities of the P- and S- waves, respectively, in the uppermost layer. Any arrivals with angles greater than θ_{crit} are rejected. For instance, a $V_p/V_s = 1.65$ yields a critical angle equal to 37°. However, for manual analysis, where the usual can identify converted phases, this can be extended to 45°, to include a larger dataset. It is noted that the selection of waveforms with clear and impulsive shear – waves through visual inspection is still crucial, given that the shear – wave window can be affected by errors and inaccuracies of the local velocity model and the hypocenter. This criterion is used to filter out the bulk of unsuitable arrivals. The implications of such a strict criterion are discussed in Chapter 5.1. Further review of eligible waveforms conducted by the analyst includes the assertion of higher amplitudes in the horizontal components than in the vertical one at the S – wave arrival, the opposite of which can be an indicator of scattering (Aster et al., 1990; Kaviris, 2003).

First, the arrival of the S_{fast} wave is identified in the horizontal components. The analyst then determines its polarization from the polarigram (or, if necessary, the hodogram). In this example, φ is equal to N97°E. This stage is prone to errors due to misidentification of the S_{fast} arrival, a result of factors such as high noise content and arrival of secondary phases.

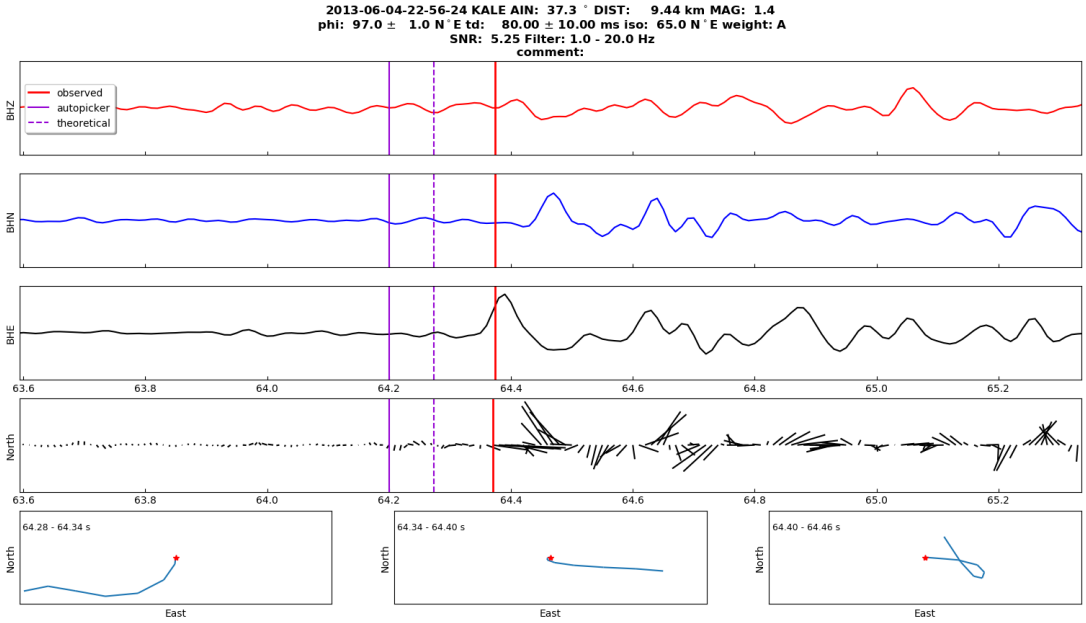

Figure 2.3. Initial stage of processing for the visual inspection of particle motion diagrams technique. Top: waveforms of the vertical, north and east components. Middle: the polarigram in the NE plane. Bottom: hodograms in the NE plane for three different timeframes, around the pick of the S_{fast} (indicated by the red vertical line). The origin time and station are presented in the figure title. "AIN": angle of incidence (°), "DIST": epicentral distance (km), "MAG": the magnitude, "phi": the polarization direction (N°E), "td": the time – delay (ms), "iso": the polarization of the corrected S – waves (N°E), "weight": a qualitative grade of the measurement, "SNR": Signal – to – Noise Ratio around the S_{fast} arrival and "Filter": the current filter limits. The dashed violet line indicates the theoretical arrival of the shear – wave and the solid violet line showcases the automatically picked shear – wave. The waveforms are filtered in the 1 – 20 Hz band.

Waveforms are then rotated to the axial system defined by the Fast and Slow polarization directions (FS), according to φ. To measure the time – delay, the S_{slow} component is temporally shifted towards the S_{fast}. The number of samples (commonly expressed in s or ms) required for the two to arrive simultaneously to the station is the t_d. In this case, the value of the time – delay is 80 ms (or 8 samples, considering that the sampling rate is 100 sps).

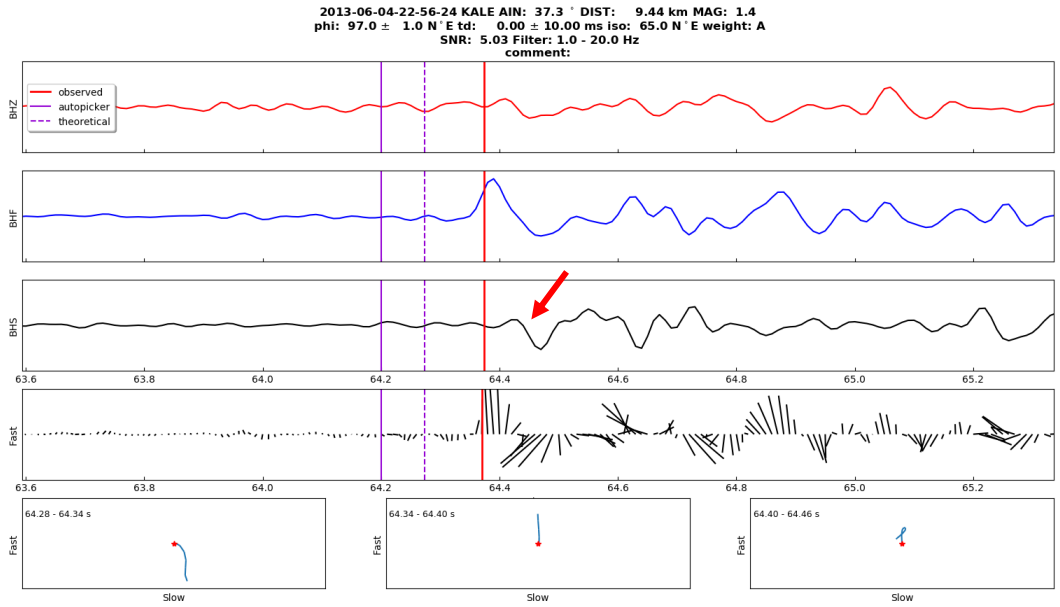

Figure 2.4. Rotated stage (FS axial system) of processing, before the correction for t_d. The red line shows the arrival of the S_{fast} and the red arrow indicates the arrival of the S_{slow}. Note that the vectors of the S_{fast} pulse are now parallel and close to F0°S (i.e. the fast direction in the FS system). Notation and filtering as in Figure 2.3.

After correcting for the time – delay, an auxiliary angle is measured between the S_{fast} direction and the displacement vector (Figure 2.5). The addition of this to φ gives the polarization of shear – waves as if the propagation medium was isotropic, this would be the S-wave polarization at the station.

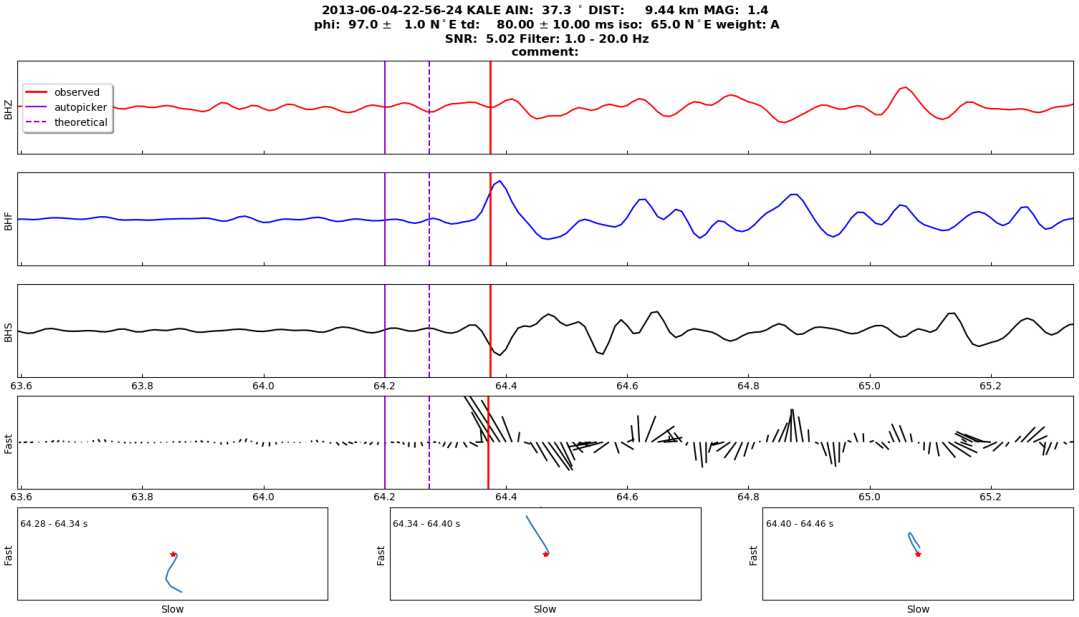

Figure 2.5. Rotated stage (FS axial system) of processing, after the correction for t_d. The parallel polarization vectors at the shear – wave arrival are used to measure the auxiliary angle. Notation and filtering as in Figure 2.3.

Finally, the waveforms are rotated back to the NE axial system (Figure 2.6). The polarization of the shear – waves is measured to verify previous results. The shear – wave should now arrive in both horizontal components at the same time. The polarization direction of the S -wave can be measured, as if the medium was isotropic (here, this is equal to N65°E). A quality weight is assigned to the measurements by the analyst. Kaviris et al. (2018a) proposed the use

of a four – stage qualitative grading system. For cases with high uncertainty in the identification of the S_{fast} and poor parallel orientation of polarization vectors with its direction, a quality weight "D" is given. Cases with undisputed arrivals and φ parallel vectors are graded "A". Intermediate situations are characterized as either "B" or "C".

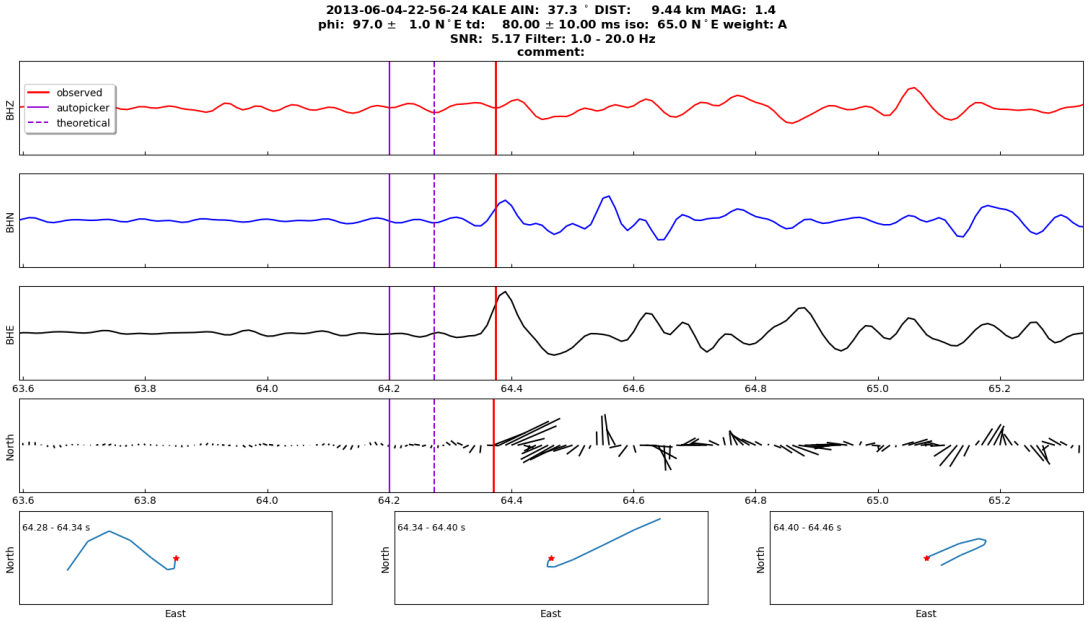

Figure 2.6. Corrected stage (NE axial system) of processing. The waveforms are the same as the ones that would have been recorded if the propagation medium was isotropic. The particle motion of the shear – wave should now be linear. The quality of the measurements is provided by the user in "weight" (in this case, the quality is "A"). Notation and filtering as in Figure 2.6.

The visual inspection of particle motion diagrams is a technique that offers the immediate engagement of the analyst with the processing and enables them to quality control every step. Nevertheless, it requires considerable human effort and time. Considering that a high amount of eligible (per the initial criteria) data is rejected after the analyst has inspected them (and, thus, has already dedicated time to the processing), the effort invested can yield a small number of results. Another major issue with any manual method is the analyst's bias. Several factors can contribute to human error (e.g. fatigue, poor psychological condition, outcome bias), hence the above method is prone to them. In addition, there is no robust quantitative error estimation in the manual analysis. The analyst can assign a qualitative weight as shown above, but this is not based on objective criteria and no comprehensive statistical analysis can be conducted.

2.2. Rotation – Correlation

The Rotation – Correlation (RC) method of SWS analysis was first employed by Fukao (1984) and Bowman and Ando (1987). Since then, it has been widely applied in several occasions of both teleseismic (e.g. Evangelidis, 2017; Evangelidis et al., 2011) and local (e.g. Bianco and Zaccarelli, 2009; Giannopoulos et al., 2015; Sharma et al., 2017) studies and has been integrated in existing software packages (Piccinini et al., 2013; Wüstefeld et al., 2008). The concept of this method is rather simplistic; horizontal waveforms are rotated to all possible S_{fast} polarization values and the Cross – Correlation Function (CCF) is obtained for a range of time – delays. The pair of parameters that maximizes the CCF (Figure 2.7) is the accepted one. The suitability of the measurement is then verified by inspecting the linearity of the particle motion for the corrected horizontal components. The RC method assumes that after removing the splitting effect, pulses of shear – waves in the horizontal components are similar.

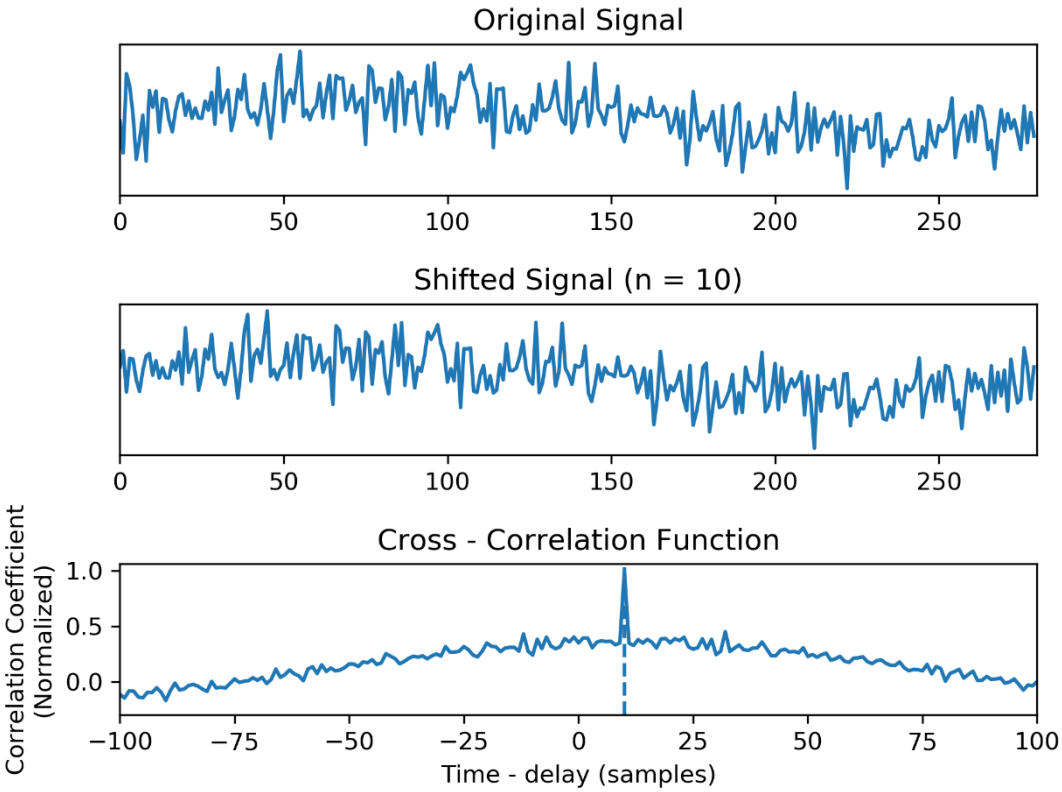

Figure 2.7. Cross – correlation of two synthetic sinusoidal wave signals. The original signal (top) is shifted by 10 samples (middle). The CCF (bottom) exhibits the expected peak of the correlation coefficient at a time – delay of 10 samples (dashed line). The shape and symmetry of the CCF is greatly affected by the random noise in the data.

At this point, it is important to define the terms "rotation" and "correlation", in the mathematical sense. A two – dimensional rotation matrix (R_{2D}) can be utilized to perform the rotation of the horizontal components to a system defined by the backazimuth. This matrix can be defined as:

$$R_{2D} = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix} \tag{2.5}$$

where θ the rotation angle. For the RC method, this is the polarization direction of the S_{fast}. The final signals of the two horizontal components (N and E) in the rotated S_{fast} (F) and S_{slow} (S) directions are obtained from:

$$\begin{bmatrix} F \\ S \end{bmatrix} = R_{2D} \begin{bmatrix} N \\ E \end{bmatrix} \tag{2.6}$$

Rotations can also be performed in the three – dimensional space. To rotate the three waveforms (ZNE) to a coordinate system defined by the ray, the following transformation needs to be applied (Plesinger et al., 1986):

$$R_{3D} = \begin{bmatrix} \cos\theta & \sin i_0 \sin\theta & \sin i_0 \cos\theta \\ \sin i_0 & \cos i_0 \sin\theta & \cos i_0 \cos\theta \\ 0 & \cos\theta & \sin\theta \end{bmatrix} \tag{2.7}$$

where θ the backazimuth and i_0 the incidence angle of the ray. The final waveforms in the ray system (LQT) are obtained by:

$$\begin{bmatrix} L \\ Q \\ T \end{bmatrix} = R_{3D} \begin{bmatrix} Z \\ N \\ E \end{bmatrix} \tag{2.8}$$

In any case, the rotation to the initial system can be achieved by multiplying with the transposed rotation matrix.

The CCF ($c(t)$) of the two signals is defined as (Shearer, 2009):

$$c(\tau) = F(\tau) \star S(\tau) = \int_{-\infty}^{+\infty} F(t-\tau) S(\tau) d\tau \tag{2.9}$$

where t the time, τ the time – delay and $\star$ the operator of the cross – correlation. It is noted that cross – correlation is not commutative, i.e. $F(t) \star S(t) \neq S(t) \star F(t)$. It is also useful to compare this operation to convolution (which could be more comprehensible in programming terms):

$$F(t) * S(t) = F(-t) \star S(t) \tag{2.10}$$

The RC method does not require extensive computer resources and is relatively simple to program, given that most modern languages contain subroutines for cross – correlation. One major issue with this processing scheme is the definition of the signal window used in the analysis. The sole user involvement in the process is to determine the time interval of the original waveforms that will be used, around the arrival of the S_{fast}. This particular problem is common in semi – automatic methods and will be discussed extensively in Chapter 2.4. Yet another issue with this technique is the complexity of the pulse. Especially in local events, the shape of the waveform can be intricate due to noise in frequencies retained for the analysis, resulting in weak correlations. Applying a strict filter will simplify the pulse but will also lead to band – limited data and possible cycle skipping effects. There is a fine balance between preserving as much of the original shear – wave energy content as possible and removing noise. Crampin and Gao (2006) pointed out that the RC method is not sensitive to the subtle changes in the signal, caused by SWS, resulting to the correlation being dominated by the large amplitudes of shear- and, even, surface- waves.

Finally, the RC method does not contain an inherent method of error estimation. Wüstefeld et al. (2008) adopted a Fisher transformation technique, according to which the normalized correlation coefficient r is transformed into the parameter z from the relation:

$$z = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right) \tag{2.11}$$

The distribution of z is similar to a normal one with a standard deviation of $\sigma_z = \sqrt{\frac{1}{n} - 3}$, where n the number of samples. For the confidence level $\mu_z = 2\sigma_z$, the confidence level of the correlation coefficient can be acquired from:

$$\mu_r = \tan(\mu_z) \tag{2.12}$$

This value is used to obtain the error bounds of φ and t_d (in a manner identical to that described in Chapter 2.3).

2.3. Eigenvalue and Minimum Energy

The Eigenvalue (EV) and, its special case, Minimum Energy (ME) methods[1] were introduced by Silver and Chan (1991). This technique is more sophisticated than the RC, but, similarly, enables the semi – automatic analysis of SWS. The EV has been extensively employed in seismic anisotropy studies in the literature (e.g. Savage, and Silver, 1993; Peng and Ben-Zion, 2004; Teanby et al., 2004; Baccheschi et al., 2008; Evangelidis et al., 2011; Walpole et al., 2014; Evangelidis, 2017; Nolte et al., 2017) and has been the prominent feature of several released software (Wüstefeld et al., 2008; Savage et al., 2010; Reiss and Rümpker, 2017). A concise summary of the method and its theoretical background, as proposed by Silver and Chan (1991), is presented below.

The concept of the EV technique assumes that SWS can be expressed as the geometric projection of the real unit vector $\hat{\boldsymbol{p}}$ (with the displacement direction) to the directions of the S_{fast} and S_{slow} (defined by $\hat{\boldsymbol{f}}$ and $\hat{\mathbf{s}}$, respectively). The splitting operator $\boldsymbol{\Gamma}$ is given by the relation:

$$\boldsymbol{\Gamma} \cong e^{i\omega\frac{t_d}{2}\hat{\boldsymbol{f}}\hat{\boldsymbol{f}}} + e^{-i\omega\frac{t_d}{2}\hat{\mathbf{s}}\hat{\mathbf{s}}} \tag{2.13}$$

In essence, the time – delay is applied in both the S_{fast} and S_{slow} components. This is why t_d is halved in the exponent. The negative sign in the part of Eq. 2.13 that refers to the S_{slow} declares that this component is temporally shifted in the opposite direction of the S_{fast}. The split waveform u_s is described by:

$$u_s(\omega) = w(\omega)e^{-i\omega T_0}\boldsymbol{\Gamma}(\varphi, t_d)\hat{\boldsymbol{p}} \tag{2.14}$$

where $w(\omega)$ the wavelet function, T_0 the arrival time of the shear – wave at the surface and φ the angle between the displacement and S_{fast} directions. After defining $\delta\boldsymbol{T} = \frac{\boldsymbol{t_d}}{\mathbf{2}}\left(\hat{\boldsymbol{f}}\hat{\boldsymbol{f}} - \hat{\mathbf{s}}\hat{\mathbf{s}}\right)$, the compact form of the splitting operator can be obtained:

$$\boldsymbol{\Gamma} = e^{i\omega\delta\boldsymbol{T}(\varphi, t_d)} \tag{2.15}$$

It is evident that if $\boldsymbol{\Gamma}$ =1, Eq. 2.14 describes a waveform that is not split, as it corresponds to the isotropic case. Thus, by searching for the inverse $\boldsymbol{\Gamma^{-1}}$, the values for φ and t_d can be determined. This can be achieved by obtaining the covariance matrix of particle motion in the S_{fast} plane in the time – domain. Its eigenvalues can be used to estimate the linearity of the motion. The covariance matrix is defined as:

$$c_{ij}(\varphi, t_d) = \int_{-\infty}^{+\infty} u_i(t)u_j(t - t_d)dt \tag{2.16}$$

where i, j the directional indices. If the propagation medium is anisotropic, c_{ij} has one non – zero eigenvalue $\lambda_1 = \int_{-\infty}^{+\infty} w(t)^2 dt$ and eigenvector $\hat{\boldsymbol{p}}$. The inverse operator $\boldsymbol{\Gamma^{-1}}$ that corresponds to a singular covariance matrix for the corrected waveform is the accepted one. However, noise will not permit the singularity and, as a result, the most nearly singular matrix is sought. To assess the linearity of the motion, Silver and Chan (1991) suggested to seek the minimum λ_2 value, which can provide insight about the variance of the noise process. In

[1] Hereafter, the term "EV" will refer to both, unless otherwise stated.

conclusion, the EV method (not its ME sub – case) uses a grid of φ and t_d pairs for which the covariance matrix of the corrected for anisotropy waveforms is obtained. The parameter pair that corresponds to the minimum λ_2 of the matrix is the accepted one. Eq. 2.13 – 2.16 refer to waveforms transformed to the LQT axial system, i.e. rotated according to the backazimuth and angle of incidence (similar to $\hat{\boldsymbol{p}}$). For teleseismic phases, where the displacement vector is known, the energy in the transverse component can be minimized instead. This modification refers to the ME method.

Silver and Chan (1991) also offered a comprehensive framework for error estimation in their method, a significant advantage over other techniques. Assuming that the noise process presents a χ^2 distribution, the confidence area a can be defined by values of λ_2^{min} that satisfy the following:

$$\frac{\lambda_2}{\lambda_2^{min}} \leq 1 + \frac{k}{\nu - k} f_{k,\nu-k}(1-a) \tag{2.17}$$

where ν the degrees of freedom, k the number of parameters (for SWS analysis these are two, hence $k = 2$) and f the inverse of the F – distribution. For a confidence level of 95%, $a = 0.05$. The estimation of ν is rather intricate and involves the spectral analysis of the waveforms. For the corrected transverse trace, the following quantities are calculated:

$$\begin{aligned} E &= \sum_{n=1}^{N-1} |f_n|^2 |g_n|^2 + \frac{1}{2}(|f_0|^2|g_0|^2 + |f_N|^2|g_N|^2) \\ E_4 &= \sum_{n=1}^{N-1} |f_n|^4 |g_n|^4 + \frac{1}{3}(|f_0|^4|g_0|^4 + |f_N|^4|g_N|^4) \end{aligned} \tag{2.18}$$

where g a Gaussian noise process, f a filter, n the sample and N the number of samples. After obtaining E and E_4, the degrees of freedom are acquired:

$$\nu = 2\left(\frac{2E^2}{E_4} - 1\right) \tag{2.19}$$

Having all the necessary elements, the 95% confidence region is defined by Eq. 2.17. To finally estimate the errors of φ and t_d, three possible cases can be distinguished. For a symmetric elliptical confidence region (Figure 2.8), the bounding values of both parameters along the two axes correspond to 2σ. For asymmetric (but ellipsoid) regions, the largest symmetric one (which

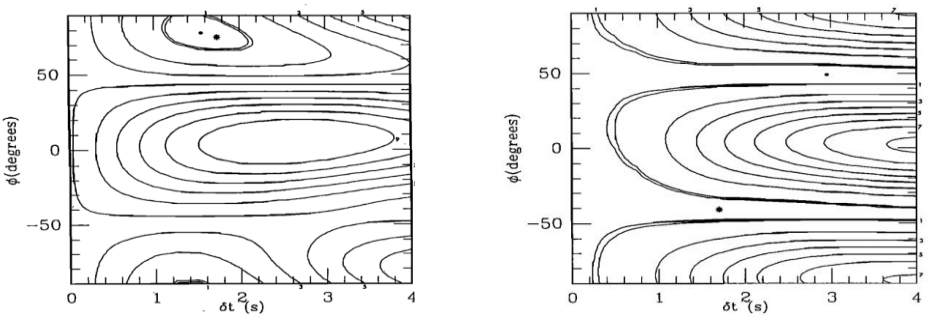

Figure 2.8. Examples of contour plots corresponding to the distribution of energy related to the EV method, from Silver and Chan (1991). Left: The accepted φ - t_d pair has coordinates N75°E and 1.75 s. The ellipsoid contours are indicative of SWS and permit the error estimation. Right: Measurement at coordinates N-39°E and 1.75s. Contours are not symmetric or elliptical, due to the absence of SWS. Double contours denote the 95% confidence region.

contains the asymmetric) is used to find 2σ. Finally, if the confidence area does not showcase any ellipticity, splitting has not occurred (Figure 2.8). The errors for each parameter are acquired by dividing the distance within the contour along each axis and dividing by four.

The EV method, albeit popular, is perturbed by several issues. As in the case of the RC, this method is sensitive to the signal window selected by the analyst (or automatically, as seen in Chapter 2.4). Even though it could be argued that it is more resource demanding than the RC, the difference should be negligible in modern computer systems. According to Crampin and Gao (2006), the EV method is not suitable for analysis of local events from surface recorders. They attributed this to the variability of local shear – wave arrivals due to two factors; S – waves being part of the larger shear- and surface- wave coda and distortions of polarization due to irregular topography. Walsh et al. (2013) argued that the error analysis procedure of Silver and Chan (1991) led to systematic underestimation of uncertainties, due to an overestimation of ν and noise not actually fitting a Gaussian distribution. To rectify the former, they proposed the following relation:

$$E_4 = \sum_{n=1}^{N-1} \frac{4}{3} |f_n|^4 |g_n|^4 + \frac{1}{3} (|f_0|^4 |g_0|^4 + |f_N|^4 |g_N|^4) \tag{2.20}$$

In practice, the coefficient of individual parts of the sum was changed from 1 to 4/3. Walsh et al. (2013) observed that by using Eq. 2.20, the overestimation of ν was reduced from a factor of 1.33 to 1.10.

2.4. Cluster Analysis in Shear – Wave Splitting

Cluster Analysis (CA) is a statistical procedure used in pattern recognition to identify groups amongst large datasets. It attempts to form assemblies of individual observations (clusters). CA is prevalent today because of its usefulness in unsupervised machine learning, i.e. in solving problems concerning the organization of unlabeled data. Essentially, it is a procedure that not only recognizes similarities among given data, but also identifies these similarities. CA has been used in Seismology to identify spatiotemporal event clusters (Shearer et al., 2005; Kapetanidis et al., 2015), classify strong – motion recordings based on their spectral properties (Ding et al., 2018) and effectively pick teleseismic phases (Houser et al., 2008). In SWS, CA has been used to resolve the conundrum of signal window selection in semi – automatic methods (as seen in Chapters 2.2 and 2.3). Window selection is the only source of human bias in both RC and EV. Especially when crustal studies are concerned, the strong shear- and surface-wavetrains could interfere with picking the end point of the analysis window properly and, as a result, lead to the contamination of the analyzed signal (Crampin and Gao, 2006). Furthermore, this process can become time – consuming, if the analyst decides to test several windows, seeking the optimal. Teanby et al. (2004) proposed a simple, brute – force (but, at the same time, elegant) solution; automatically perform multiple SWS measurements over a range of predefined windows and use CA to select the best. Their method has been employed in the

mFast software (Savage et al., 2010). Following, the fundamental principles of the method will be summarized, as first published by Teanby et al. (2004)[2].

First, the range of the predefined signal windows must be configured. The start (T_{beg}) and end (T_{end}) points, in relation to the S_{fast} arrival, are combinations of the following calculated times:

$$\begin{aligned} T_{beg} &= T_{beg1} - (i-1)\Delta T_{beg} \\ T_{end} &= T_{end0} + (j-1)\Delta T_{end} \end{aligned} \tag{2.21}$$

where T_{beg1} the maximum accepted value for T_{beg}, ΔT_{beg} the time interval for the start point definition, i in the range $[1, N_{beg}]$ (with N_{beg} being the number of start points), T_{end0} the minimum accepted value for T_{end}, ΔT_{end} the time interval for the end point definition and j in the range $[1, N_{end}]$ (with N_{end} being the number of end points). The boundaries of the accepted time points should be defined to include a sufficient, for the SWS analysis, part of the S – wave energy, while, at the same time, not being contaminated by high noise (e.g. the arrival of P – waves). For example, T_{beg1} and T_{end0} that permit a 5 min window around the shear – waves would be unsuitable, given that it would contain P – waves, as well as the shear- and surface-wave codae. It is evident that the maximum number of windows is $N = N_{beg}N_{end}$. For each window, a semi – automatic method is applied (Figure 2.9). Teanby et al. (2004) used the EV technique, but in the Pytheas software RC can also be used.

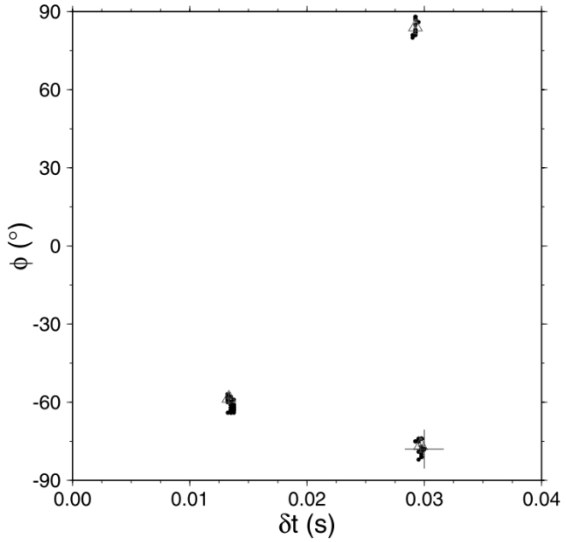

Figure 2.9. An example of the distribution of measurements in 250 time windows (Teanby et al., 2004), where CA will select the optimal cluster and measurement pair (in this case, marked with the cross).

At this point, there are N measurements, possibly comprising individual subsets of identical results. Initially, the obtained parameters need to be normalized, to be comparable. If this procedure is not conducted, the much larger (numerically) values of φ will exhibit higher weights in the linkage process than the t_d. The fourth standardization of Milligan and Cooper (1988) is used:

[2]The topic of Cluster Analysis may seem exotic to several Seismologists and Geophysicists. For this reason, the description of the method in the current Thesis is more extensive, than the previous ones (PM, RC and EV). Still, Teanby et al. (2004) offer a much richer explanation of each segment of the method, as well as the full references. Their article is a highly suggested read by the author.

$$Z_4 = \frac{X}{X_{max} - X_{min}} \tag{2.22}$$

where Z_4 the standardized value, X the individual measurements and $X_{max} - X_{min}$ their range. The data are now ready to be clustered. To achieve this, an unsupervised hierarchical agglomerative clustering procedure is used (Everitt et al., 2011). Starting with N singletons (clusters containing a single observation), $N - 1$ linkages are performed to end up with a cluster containing N observations. At each stage, the intercluster distances are calculated between cluster centers. This is the Euclidian distance:

$$d_{kl} = \sqrt{(T_{dk} - T_{dl})^2 + (\Phi_k - \Phi_l)^2} \tag{2.23}$$

where T_d and Φ the time – delay and polarization direction values corresponding to the centers of the clusters k and l. For each cluster j, the center coordinates are defined as the means of the individual observations i that belong to j:

$$\begin{aligned} T_{dj} &= \frac{\sum_{i=1}^{N_j} t_{di}^{(j)}}{N_j} \\ \Phi_j &= \frac{\sum_{i=1}^{N_j} \varphi_i^{(j)}}{N_j} \end{aligned} \tag{2.24}$$

where N_j the number of data points in the cluster. For descending number of clusters M belonging in the $[N, 1]$ range, the two clusters with the minimum distance are combined, in every iteration. In this way, a hierarchy is obtained, with formed clusters for different values of M. To determine the optimal M, two criteria are used. The Caliński and Harabasz (1974) criterion calculates a critical value $c(M)$ for each number of clusters:

$$c(M) = \frac{(N - M)trace(\boldsymbol{B})}{(M - 1)trace(\boldsymbol{W})} \tag{2.25}$$

where $\boldsymbol{B}$ and $\boldsymbol{W}$ the between- and within- cluster covariances, respectively, as defined in Teanby et al. (2004). The value of M that maximizes c is considered the optimal. A high c value indicates greater intercluster distances and relatively tight cluster formations. The second criterion is that of Duda and Hart (1973). The null hypothesis (i.e. the two clusters being combined as a single cluster) is rejected when:

$$\sqrt{\left(1 - \frac{\sigma_2^2}{\sigma_1^2} - \frac{2}{\pi p}\right)\left(\frac{N_j p}{2\left(1 - \frac{8}{\pi^2 p}\right)}\right)} > c_{critical} \tag{2.26}$$

where σ_2 and σ_1 the variances of two and combined clusters, respectively, as defined in Teanby et al. (2004), p the number of variables (i.e. two) and $c_{critical}$ a threshold value equal to 3.2. This criterion is applied by increasing the number of clusters at every iteration (from 1 to N), until Eq. 2.24 is not satisfied. As a final safeguard against erroneous selection of M, an upper limit (M_{max}) of its acceptable values is predefined by the analyst. Finally, the optimal cluster must be selected, out of the M specified clusters. An initial filtering is conducted, by rejecting

all clusters with a number of observations less than $N_{c_{min}}$. To select the final cluster, the within – cluster ($\sigma_{c_j}^2$) and mean data ($\sigma_{d_j}^2$) variances are calculated:

$$\sigma_{cj}^2 = \frac{\sum_{i=1}^{N_j}\left(t_{di}^{(j)} - T_{dj}\right)^2 + \left(\varphi_i^{(j)} - \Phi_j\right)^2}{N_j}$$
$$\sigma_{dj}^2 = \left[\sum_{i=1}^{N_j}\frac{1}{\left(\sigma_{t_{di}}^{(j)}\right)^2}\right]^{-1} + \left[\sum_{i=1}^{N_j}\frac{1}{\left(\sigma_{\varphi_i}^{(j)}\right)^2}\right]^{-1} \quad (2.27)$$

where σ the uncertainties acquired for each parameter from the respective error estimation procedure of the used SWS analysis method. The maximum of the two variances obtained from Eq. 2.27 is the overall cluster variance $\sigma_{o_j}^2$. The optimal cluster (and, consequently, pair of parameters) is the one with the smallest overall variance.

CA is a very intriguing topic in SWS, with a lot of potential. Nevertheless, it is currently held back by inaccuracies in the existing underlying methods. No matter how accurate the clustering is, systemic errors in the determination of SWS parameters are still the issue. Another issue is the efficiency of the method in terms of production. While it is true that after the initial configuration there is no need for human interaction, the analyst must take great care in selecting the parameters, especially the number and interval of time windows. Savage et al. (2010) proposed a modification to the above process that could resolve some of those issues. Before the SWS analysis starts, waveforms are filtered in various bands to determine the best one, based on the SNR. The predefined time windows are then selected from the period and pick of the shear – waves. By implementing those additional steps in the processing scheme, they observed a general agreement of CA results with measurements obtained manually.

2.5. Shear – Wave Splitting Tomography

Values of time – delays are commonly obtained for single points in space and are interpreted as such. However, even in areas covered by dense local networks, their spatial distribution cannot be inferred. Tomographic inversion methods have been previously used to obtain both vertical and lateral variations of seismic velocities, to infer the underlying structure (e.g. Voulgaris, 1991; Drakatos et al., 2005; Karakonstantis, 2017). The same principle can be applied on time – delays acquired from SWS analysis. Zhang et al. (2007) used a 3D shear – wave velocity model to project SWS time – delays along the obtained ray paths in the vicinity of the San Andreas Fault, California, USA. Their results exhibited some correlation with potential sources of anisotropy.

Johnson et al. (2011) developed a method (Figure 2.10) that does not require the rigorous task of acquiring a S – wave tomography first. This concept is based on two assumptions: (a) the t_d is accumulated along the ray path and (b) the total time – delay is the sum of t_d for each grid block that the ray crosses. The measured time – delay of the ray (t_{d_r}) is:

$$t_{d_r} = \sum_{b=1}^{n} s_b L_{rb} \tag{2.26}$$

where s_b the strength of anisotropy and L_{rb} the length of the ray path in each grid block b. To isolate areas with sparse ray paths, a quad tree gridding is employed, where each node is separated in four new ones, until the maximum number of rays is met. The problem to solve for the tomographic inversion is:

$$\boldsymbol{G}^T \boldsymbol{C}^{-1} d = (\boldsymbol{G}^T \boldsymbol{C}^{-1} \boldsymbol{G}) m \tag{2.27}$$

where $\boldsymbol{G}$ the design matrix containing all L_{rb} values, d the t_d values for each ray, m the model containing the strength of anisotropy per block and $\boldsymbol{C}$ the error covariance matrix defined as:

$$\begin{aligned} C_{ii} &= \sigma_d^2(i) + \frac{L_b^2 \sigma_m^2 n_b(i)}{n} \\ C_{ij} &= \frac{\sigma_m^2 L^2}{n^2} \end{aligned} \tag{2.28}$$

where σ_d^2 the squared standard error of the ith measurement, n_b the number of grid cells that the ray intersects and σ_m^2 the mean of the variance of the result. Checkerboard tests are conducted by obtaining theoretical t_d values from:

$$d_{CB} = \boldsymbol{G}_{CB} m_{CB} \tag{2.29}$$

where CB denotes the checkerboard test. In addition to the tomographic inversion for the time – delays, Johnson et al. (2011) developed a weighted spatial averaging for φ values. Polarization directions are averaged in each block, based on neighboring values. The weighting is conducted through one of the following methods: (a) no weighting, (b) dividing by the distance, (c) dividing by the squared distance and (d) by taking into account the results of the tomography. In the last case, the weighted mean polarization direction ($\overline{\varphi}_b$) is:

$$\overline{\varphi}_b = \tan^{-1} 2 \left(\frac{\sum_{r=1}^{n} \mathrm{w_{rb}} \sin \varphi_r}{\sum_{r=1}^{n} w_{rb}}, \frac{\sum_{r=1}^{n} \mathrm{w_{rb}} \cos \varphi_r}{\sum_{r=1}^{n} w_{rb}} \right) \tag{2.30}$$

where $w_{rb} = \frac{s_b}{t_{d_r}}$. The above methods are included in the Tomography Estimation and Shear-wave-splitting Spatial Average (TESSA) software (Johnson et al., 2011).

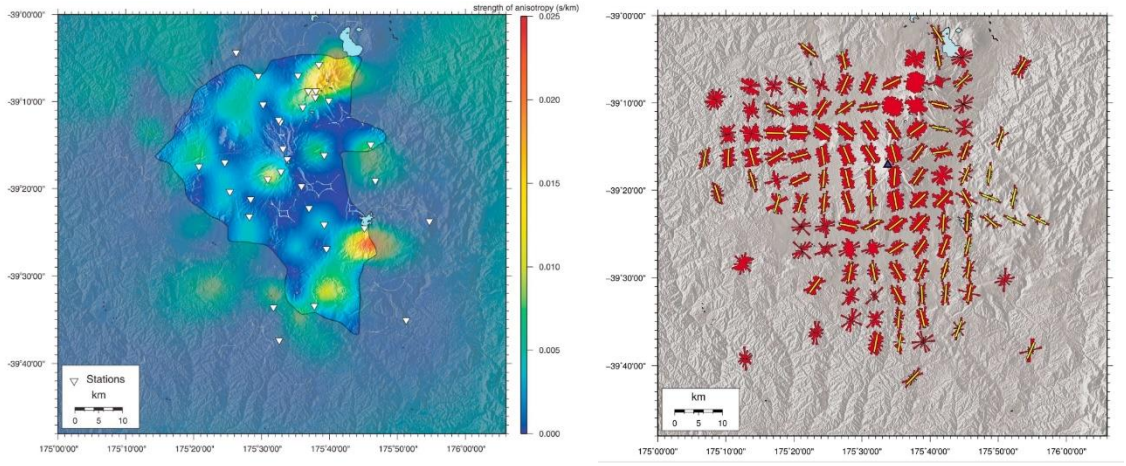

Figure 2.10. Results of SWST (left) and spatial averaging of φ (right) from Mt. Ruapehu, New Zealand (Johnson et al. (2011). Two distinct seismic anisotropy regimes were identified: one controlled by stress and one by structural features.

The SWST technique was also employed in the case of the Tongarino Volcanic Center in New Zealand (Johnson and Savage, 2012), where a spatiotemporal connection was identified between an area of strong anisotropy and a contemporary eruption. In tectonic environments, SWST led to the linkage between intense anisotropy and local geological structures in Indonesia (Syuhada et al., 2017). Tomographic inversion of t_d could prove to be a valuable tool in the future. Assuming that concentrations of seismic anisotropy are connected to areas of accumulated stress, SWST would find applications in several fields, from volcanic monitoring to oil exploration and production. In active areas, the installation of a dense seismic network would even permit the detailed observation of temporal variation of anisotropy. It is important to mention at this point that seismic ambient noise, induced by ocean – related phenomena (Spingos et al., 2018) has also been used to investigate spatial variations of seismic anisotropy (Das and Rai, 2017).

3. The Pytheas Software

There have been several programs concerning SWS analysis released in the past decade. Splitlab (Wüstefeld et al., 2008) is probably the most popular, as it has been used by multiple research teams in literature (e.g. Diaz et al., 2010; Evangelidis, 2017; Kaviris et al., 2018c). It is written in the MATLAB programming language (The Mathworks Inc., 2018) and offers functionality, in terms of cataloguing, acquiring station metadata and plotting of results, but its main feature is the use of all semi – automatic methods (RC, EV and ME) concurrently in a user – friendly GUI. It also includes an interface for the TauP toolkit (Crotwell et al., 1999), which permits the calculation of theoretical arrival times for the phases to be processed (usually teleseismic). mFast (Savage et al., 2010) uses the methods of Teanby et al. (2004) and Silver and Chan (1991) in its processing scheme, incorporating the error estimation corrections of Walsh et al. (2013). The core of the software package is written in Fortran and the analyst must use BASH scripts, hence, rendering the use of a UNIX Operating System (OS) mandatory, unless the user is willing to go to great lengths to emulate such an environment in Windows (something that is not always guaranteed to work). mFast is reliant on the SAC (Helffrich et al., 2013) and GMT (Wessel and Smith, 1991) packages, as well as TauP. The prior knowledge of those additional codes, the absence of a GUI and the requirement of using UNIX could pose a significant challenge to analysts not familiar with the above. SplitRacer (Reiss and Rümpker, 2017) is also developed in MATLAB and utilizes the ME method. This software focuses on secondary features, i.e. data acquisition and event selection, but also incorporates tools for interpreting SWS results jointly, by fitting either one or two anisotropic layers. SplitRacer features a GUI, but is limited in the analysis of teleseismic events. Finally, ANISOMAT+ (Piccinini et al., 2013) uses several steps in checking the eligibility of provided data, before applying the RC method. The software is comprised by MATLAB scripts.

The above are some examples of published codes for SWS analysis. Nevertheless, there are several issues that need addressing. First and foremost, MATLAB is, in essence, a proprietary software requiring a substantial investment to acquire (although, several research institutions provide licensing options to their members). Furthermore, the initial step of archiving the data in the format that the software mandates could prove difficult and time – consuming. The requirements of specific data formats (e.g. SAC only) and the embedment of event and station metadata in the waveform files (e.g. event coordinates) are two examples of the hurdles the analyst could meet. Lastly, OS exclusivity and installation of several prerequisites from various third – party sources do not provide an optimal user experience.

3.1. Python in Science

Python (van Rossum, 1995) is an active general purpose programming language. It focuses on code readability, which enhances its ease of use by both amateur and experienced programmers. Python is growingly becoming one of the predominant languages in Science, being taught in undergraduate and postgraduate curricula of multiple fields (e.g. Biology, Neuroscience, Geophysics) in universities around the world. The libraries Scipy and Numpy (Oliphant, 2007) constitute the core of scientific computing in Python. Matplotlib (Hunter, 2007) offers a reliable and comprehensive way to create figures. Furthermore, the Obspy (Beyreuther et al., 2010; Krischer et al., 2015) library is a popular toolkit for seismology and Scikit – learn (Pedregosa et al., 2011) provides several easy – to – implement machine learning routines. The Qt5 (The Qt Company, 2018) framework, as well as its Python bindings included in PyQt5 (Riverbank Computing, 2018), can be used to develop a fully functional GUI. The above packages are described in further detail in the following chapters. Finally, it is worth mentioning that developing any Python software without interactively testing the code in iPython (Pérez and Granger, 2007) is virtually unthinkable. Any Python package mentioned in the current thesis can be installed easily through the PyPi (https://pypi.org/) repository and its accompanying application (pip).

3.1.1. SciPy and NumPy

The basis of any numerical operation in Python is conducted through the SciPy and NumPy libraries. These include several modules for signal processing, statistical analysis and linear algebra, among others. NumPy contains the definition for the basis of efficient computations in Python, i.e. array objects of N dimensions. Lists (the *de facto* data structure) contain pointers to objects. In contrast, arrays have the benefit of the locality of reference, i.e. the storage of individual elements close in the memory. Furthermore, any operation between arrays is conducted through the optimized underlying code of NumPy (written in C). For example, the addition of two arrays will not be performed by iterating over each element and then adding them, but, instead, a Central Processor Unit (CPU) vector operation will be used (where instructions for the processor are executed over groups of multiple values at once). NumPy is used for general tasks, while SciPy comprises of more specialized functions, although there is some overlap.

3.1.2. Matplotlib

Matplotlib is the most popular 2D graphics package for Python in Science. It can be used to easily generate figures of high quality and embeddable to any GUI. It also enabled advanced plotting features, such as LaTeX functionality (to permit the use of scientific equations and notation in figures) and contouring. One of the main benefits of Matplotlib is its capability of interactive plotting. It includes a robust event handling system, where user interaction (i.e. mouse clicks, wheel use, key pressing and key release) can be registered and redirected to an underlying function. Moreover, users can select visually any plotted object (e.g. lines or polygons). It also offers more specific functionality through existing toolkits, such as 3D plotting and cartography. Nevertheless, the library can have an impact on performance, especially in older systems, when working with real – time plots.

3.1.3. Obspy

Obspy is a package that aided in increasing the popularity of Python in Seismology. It comprises of a multitude of modules for basic and advanced tasks. It offers a very convenient environment for data management. Most popular waveform data formats are supported, either for reading or writing (e.g. GSE2, MSEED, SAC, SEGY), rendering conversions trivial. Another significant feature of the package is the metadata handing. Information about stations can be accessed and stored in formats such as StationXML (https://www.fdsn.org/xml/station/) and SEED. Concerning event and catalogue data, the supported formats include QuakeML (Schorlemmer et al., 2011) and the CMT format (Ekström et al., 2012). Obspy offers clients for communication with several services, most importantly web services of the international Federation of Digital Seismograph Networks (FDSN) and the Incorporated Research Institutions for Seismology (IRIS). The functionality of the library (which is still getting updates) extends to tasks such as cross – correlation implementation to work with Obspy objects, spectral analysis processes and auto – picker routines.

3.1.4. Scikit – learn

This library is comprised of machine learning algorithms. The simplicity of the implementation of this package encourages researchers that do not specialize in such complicated problems to easily use them. Scikit – learn includes routines for both supervised and unsupervised learning, incorporating methods such as hierarchical clustering. The package achieves efficiency by using compiled code.

3.1.5. Qt5

Qt5 is the latest (as of the writing of this Thesis) release of the Qt framework. It permits the creation of GUIs and applications, either through the Designer tool or purely programmatically. While the library is written in C++, it is integrated into the Python ecosystem through dedicated bindings (PyQt5). It is cross – platform (permitting its use in both Windows and UNIX systems) and has been used in the development of popular software, such as Google Earth (https://www.google.com/earth/). Qt programs present a native (to the OS) interface with integrated widgets (e.g. toolbars and Matplotlib embedded plots). Qt5 also offers non – GUI features, e.g. a thread – managing subsystem.

3.2. General Description of the Software

The Pytheas software is a package written purely in Python and offering an intuitive GUI for the analysis of SWS. Even though it is not possible to test the software in a wide variety of systems in the context of a MSc Thesis, due to obvious limitations, its performance was deemed adequate in the systems presented in Table 3.1 by three analysts. A modern multi – core processor, at least 4 GB of memory and a dedicated graphics processor are recommended. The fully automatic analysis for the current thesis was conducted with "System C".

Table 3.1. Specifications of three testing systems used to evaluate the qualitative performance of the Pytheas software. Among three users, System C provided the best performance and System A the worst.

Component	System A	System B	System C
Processor	Intel i3 2350M (2 cores at 2.30 GHz)	AMD FX – 4170 (4 cores at 4.20 GHz)	Intel i5 8400 (6 cores at 4.00 GHz)
Memory	4 GB (DDR3 at 1066 MHz)	8 GB (DDR3 at 1600 MHz)	8 GB (DDR4 at 2666 MHz)
Graphics Processor	Intel HD Graphics 3000	MSI GTX 950	MSI GTX 950
Monitor Resolution	1600 x 1000	1920 x 1080	1920 x 1080
OS	Ubuntu 17.04	Windows 10 & OpenSUSE 42.1	Windows 10 & OpenSUSE 42.1

The GUI is designed to accommodate all the elements required for the analysis, so as to avoid any overlaps between them and provide a clear and pleasant working environment. Although the window is adjustable, a 1920 x 1080 resolution is recommended. There are six menus available to the analyst:

- File: Options to define a catalogue and a waveform master path (which contains the data at the end of the tree), save results of the analysis and close the program.
- View: Options that control the visual part of the application.
- Visual: Offers controls for the PM method. The analyst can switch between the three analysis stages and manage, if required, the quantities related to this technique.
- Splitting: Provides a selection of the semi – automatic and automatic techniques. For a given event – station pair, the RC, ME, EV and CA (in combination with either RC or EV) methods can be applied through this menu. In addition, the CA technique can be employed for a whole catalogue, without the intervention of the analyst. The shear – wave window can be set through this menu.
- Navigate: Selection of browsing the events of the provided catalogue or the available stations for the selected event.
- Tools: Secondary tools (e.g. filtering and spectrum viewing).

To begin analyzing data, the user must go through the "Open Catalogue" option in the "File" menu. They are then prompted to select a master path for the data and a catalogue file, through the OS's native file browser. Even though the exact structure of the data tree is irrelevant, the program will search for a directory (named after the origin time of each event, in any format recognized by Obspy's "*UTCDateTime*" function, e.g. 1992-09-03-19-14-03) at the end of each

sub – path found. This folder must contain the waveforms of the corresponding event. There is no required convention for naming the trace files. Any Obspy compatible format should work properly (the software has been extensively used with SAC files, one for each trace). When attempting to access an event – station pair for analysis, the program will check whether there is an issue with the corresponding data (e.g. missing a component) and will prompt the user to skip the pair, if needed. Concerning the catalogue, a QuakeML file must be provided.

On opening a new event – station pair, an automatic picking method will determine the arrivals of shear – waves (see Chapter 3.2.1), noted by a violet solid line (Figure 3.1). The theoretical arrival will also be shown, if calculated (dashed violet line). The analyst can control the view of the waveforms and the polarigram by using the mouse wheel. Simply using the wheel will adjust the amplitude of the waveform. If the Shift keyboard modifier is used, the limits of the time axis can be controlled. For the PM and CA methods the S – arrival can be picked by left clicking and using the Ctrl modifier. A solid red line showing the picked time is then drawn. In addition, the hodograms are updated to showcase the particle motion around the arrival (3 samples before and 3 after) and the SNR is calculated (see Chapter 3.2.1). The φ value corresponding to the picked vector is also acquired from the polarization vector. If the user does not wish to measure the polarization, the right mouse button can be used instead, to only pick the S – arrival. It is noted that, upon opening a new station – event pair, if the QuakeML file contains arrival information, the relevant pick will be indicated by the red solid line.

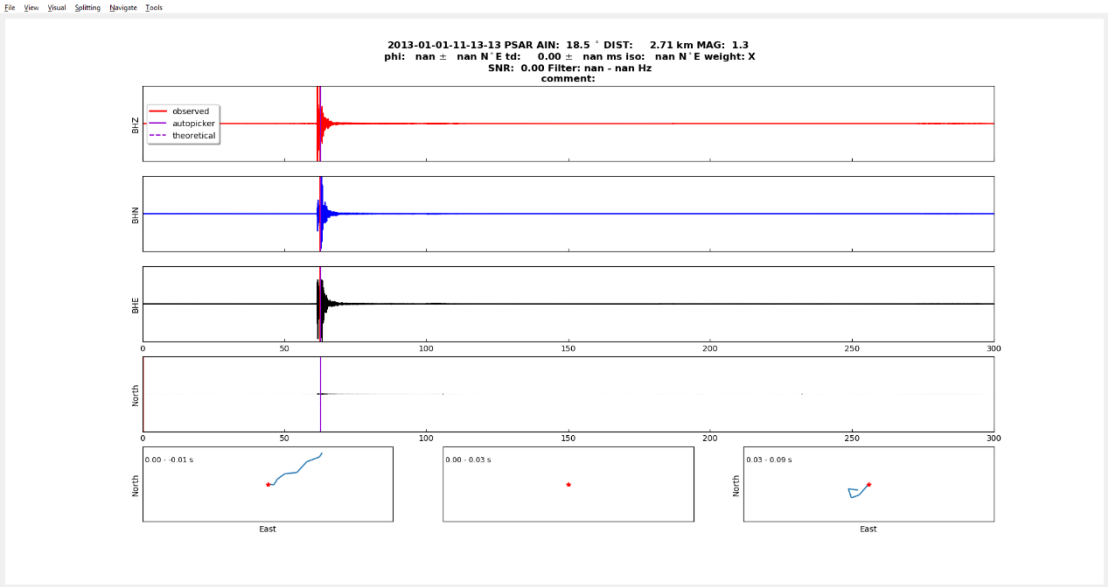

Figure 3.1. The initial screen, after the opening a new event – station pair. Values in the second and third rows of the title (which are measurement related) are placeholders. The hodograms correspond to the first 15 samples of the waveform. The various menus are located to the top left of the window.

Concerning the PM method, the waveforms can now be rotated to the FS axial system. At this stage, the S_{slow} component can be temporally moved to determine the t_d. This can be achieved either by the "Visual" menu's "Set time – delay" option (measured in ms) or by using the arrow keys. The Left arrow shifts the waveform towards the start time (positive defined time – delays) and the Right arrow moves the trace towards the end time (negative defined time – delays). Each key stroke corresponds to shifting one sample. Nevertheless, the software offers the ability of shifting at the opposite direction. At this stage, by left clicking using the Ctrl

modifier in the plot, the auxiliary angle can be acquired. The isotropic polarization will automatically be calculated. Finally, after rotating back to the NE system, the analyst can pick the arrival of the corrected S – waves to obtain the isotropic polarization (and confirm their measurement). In addition, the t_d value can be changed, so that minor corrections can be made without requiring switching between stages. A grade can also be defined at any time, through the "Visual" menu. The accepted grades are: "A", "B","C", "D" ,"E" and "N". "E" is proposed to be used for rejected measurements and "N" for null ones (see Chapter 3.3.1). The "X" grade is also available, if the user wishes to not perform any grading.

The semi – automatic methods require the selection of an analysis window. This can be achieved by using the Shift modifier and the two mouse click buttons. Left clicking sets the start of the window and right clicking the end. After selecting the method through the "Splitting" menu, a diagnostic plot is produced to evaluate the measurement. This contains part of the initial waveforms with the selected window marked, the traces in the FS system (after the temporal shift) and the final corrected data in the NE system, as well as the hodograms of the original and corrected data in the selected window. In addition, a contour plot of the relevant selection value is shown: the correlation coefficient for RC, the minimum eigenvalue for EV and the energy in the corrected transverse component for ME. The contours represented the selection value normalized per the one at the 95% confidence interval (noted by the heavier black contour). In any case, the analyst can save their results in a "spl" format file (Table 3.2) in the "splitting" directory (located in the main program's folder). These files include all necessary information related to the analysis in a space – separated array structure.

Table 3.2. Fields corresponding to the "spl" file format. If a field is not eligible (e.g. the isotropic polarization when using the CA method), the "nan" value is provided. In the Method field, the following codes correspond to the methods: "MAN" for PM, "MSC" for EV, "MME" for ME, "MRC" for RC, "ASC" for CA with EV and "ARC" for CA with RC.

Column(s)	Parameter
1 – 6	Origin Time (Year, Month, Day, Hour, Minute, Second)
7 – 8	Event Coordinates (Latitude, Longitude)
9	Depth (km)
10	Magnitude
11	Epicentral Distance (km)
12	Azimuth (N°E)
13	Angle of Incidence (°)
14	φ (N°E)
15	t_d (ms)
16	Isotropic Polarization (N°E)
17	Station Name
18	Method
19	SNR
20	σ_φ (N°E)
21	σ_{td} (ms)
22	Grade
23	Comment

If the user wishes to apply the CA method for a single event – station pair, they do so through the "Splitting" menu. At the end of the analysis, a diagnostic plot detailing the cluster selection process is displayed. Alternatively, CA can be employed for the whole dataset, as defined by the provided catalogue. The configuration file for the CA method is located in "etc/options/teanby.cnf" (for more details, see Chapter 3.3.3).

Pytheas consists of the main program and five auxiliary modules that comprise the underlying library. Following, a detailed description of each software component is presented. Functions are formatted in italics and classes in bold.

3.2.1. pytheas.py

This file contains the main class (a template containing initial values and underlying functions for the Pytheas object) and the launch code of the application. The creation of the GUI is handled by the initialization function (*__init__*). This includes the setup of the menus, assignment of GUI callback functions, creation of the initial figures and parsing of configuration files. The basic archiving and data management functions are also present. For example, *readQML* is used to parse QuakeML files, *getTree* obtains the event directories and *writeSplittingDict* is responsible for writing the analysis results to "spl" output files. *eventCorrect* matches origin times from the catalogue to event directories.

To calculate the SNR, the *calcSNR* function uses the following relation:

$$SNR = \frac{|A|_{max}}{\sigma^2_{noise}} \tag{3.1}$$

where $|A|_{max}$ the maximum absolute amplitude in the signal window, $\sigma^2_{noise} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$ with x corresponding to the amplitude of the noise and n the number of samples. This quantity is obtained for the two horizontal components and their average is the final SNR. The start of the noise window is set 1.5 s before the S – arrival and the end of the signal window 2.0 s after.

To obtain the arrival time of the S – waves automatically, the Obspy implementation of the Auto Regression – Akaike Information Criterion (AR – AIC) method by Akazawa (2004) is used. This technique is applied through the following algorithm. First, the filtered waveform (a_{v1}) is transformed according to:

$$a_{v2}(t) = \frac{|a_{v1}(t)|}{|a_{v1}|_{max}} - \frac{a_{v1}(t)^2}{a^2_{v1_{max}}} \tag{3.2}$$

Then, the envelope $a_{v3}(t)$ is acquired from:

$$\begin{aligned} a_{v3}(t) &= a_{v2}(t-1), if\ \ a_{v2}(t) \le a_{v2}(t-1) \\ a_{v3}(t) &= a_{v2}(t), if\ a_{v2}(t) > a_{v2}(t-1) \end{aligned} \tag{3.3}$$

A STA/LTA filter is applied in the window between the start of the waveform and the maximum of the envelope (i_1). The peak of the STA/LTA application defines the new end of the window (i_2). AIC is utilized to select the best fitting AR model to $a^3_{v1}(t)$, in the previous interval. The new window starts at i_2 and ends at $3i_3 - i_2$, where i_3 the point of lowest AIC. The AR – AIC method is then applied again, this time to $a^3_{v1}(t)$, but using the interval specified above. The minimum value of AIC at this stage defines the P – arrival (i_4). To estimate the S – onset, a STA – LTA filter is applied in the predominant horizontal component, between i_4 and the end of the trace. The reverse direction is also used. Finally, the AR – AIC method is applied in the unfiltered waveform of the predominant horizontal component, in the interval defined by the maximum LTA – STA value in the forward direction and the minimum STA – LTA in the reverse. The various parameters related to this method, used by Obspy, are provided in the "etc/options/picker.cnf" file (see Chapter 3.3.3)

Finally, this file contains the functions that control the secondary GUI elements. This refers to the progress bar window (*prgBar*), the event (*eventsListWindow*) and station

(*stationsListWindow)* selection windows, as well as the user prompt (*replyMsgBox*) and warning (*warnMsgBox*) windows. Also, two custom exceptions are defined: the **StationException** refers to errors occurring while navigating through station – event pairs and **EventException** when alternating between events.

3.2.2. parsers.py

This module contains classes that are used to process settings files related to various functions of the software. The module "configparser", included in the Python Standard Library (PSL), is used to read configuration files of a structure similar to that of INI files in Windows. These are composed by sections comprised of separate keys (properties corresponding to variables). Each property acquired from the configuration file is stored as a unique member variable of each class, to be used in the program. Pytheas currently uses four such files: (i) picker.cnf, containing settings about the implemented auto – picker routine (see Chapters 3.2.1 and 3.3.3), (ii) taup.cnf (see Chapter 3.2.6), (iii) clustering.cnf (see Chapter 3.2.5) and (iv) grading.cnf (see Chapter 3.3.1). As an example of the structure of those files, picker.cnf consists of one section (ARPICKER) which holds eleven keys (e.g. the frequencies for the bandpass filter used in the auto – picking process, defined as the variables f1 and f2). The parsing of the configuration files takes place in the program's initialization stage.

3.2.3. rotationcorrelation.py

This module contains the *crosscorrelation* function, used to apply the RC method. The function accepts as arguments the Obspy stream object containing the three traces, the tuple of the selected window indices and (optionally) the maximum time – delay for which to calculate the cross – correlation function (CCF), by default set at 250.0 ms. The φ and t_d test matrices are initialized at the start, as well as the array that will receive the CCFs. For each test polarization direction, the original waveforms are rotated to the FS axial system and the CCF for the maximum time – delay is obtained. After all the φ trials are conducted, the parameter pair with the maximum absolute correlation coefficient is selected as the measurement. To obtain the errors, the 95% confidence region for the coefficient is defined the 95% of its value (Piccinini et al., 2013). The error bounds are then calculated from the distance between the measurement and the contour defined by the 95% confidence level value. The output of the function is the RC results and corresponding errors, the trial matrices and the error estimation relevant information.

3.2.4. eigenvalue.py

All the functions related to the EV method are defined in this module. *SilverAndChan* is used as the main procedure. It requires the following arguments: (1) the stream containing the three waveforms to analyze, (2) the backazimuth of the station and (3) the user – defined window where the method will be applied (in samples). The latter is selected by the analyst in the application's main window or provided by the cluster analysis routines. The general algorithm used to apply this method is shown in Figure 3.2.

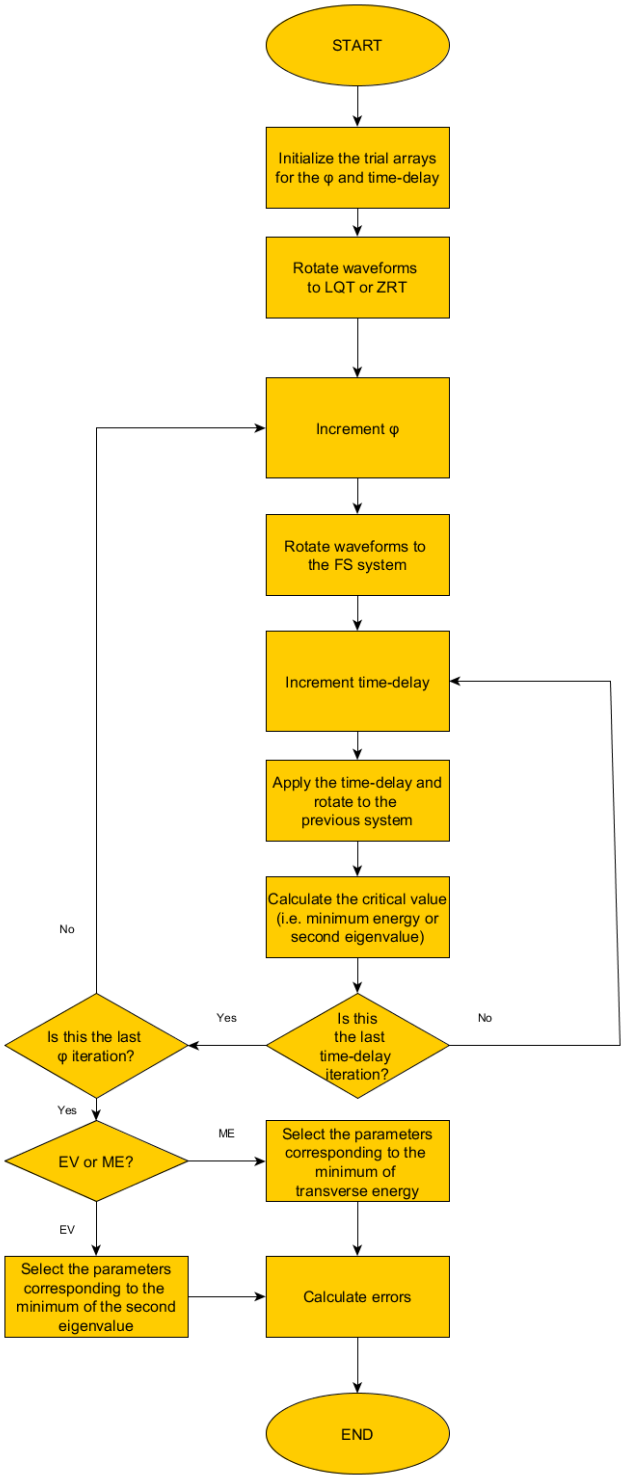

Figure 3.2. Flow chart showcasing the algorithm used in the implementation of the EV (and ME) method in the Pytheas software.

The angle of incidence can also be used. If that is the case, the waveforms will be rotated to the LQT system, instead of the ZRT. The maximum time – delay to be tested and the method used (EV or ME) can also be passed as arguments to the function (the default values are 0.250 s and EV, respectively). The trial arrays are first initialized (for increments of 1° for φ and 1 sample for t_d). The φ array contains values between -90 and 90 N°E, while the t_d trials range

from 0 to the number of samples given as argument (default is 250 samples for a 100 sps waveform). The arrays for the critical value (either minimum energy or λ_2), the two eigenvalues and the first eigenvector are initialized (starting values are zeroes). All matrices to be used are created at the start, to render the following calculations faster. The algorithm then performs the rotation of the initial ZNE waveforms to either the LQT or the ZRT system. To optimize iterating over the two test anisotropy parameters, the "itertools" module of the PSL is used. It was observed that using "itertools" led to faster calculations, compared to using a nested loop, by 0.8 s (for each iteration). For each combination of φ and t_d, the waveforms are first rotated to the FS axial system and the time – delay is applied. Then, they are rotated back to the radial system and the covariance matrix of the corrected radial and transverse components is obtained, as well as the energy in the transverse component. The eigenvalues and eigenvectors are stored in the respective arrays. After that, depending on the selected method, the parameters pair that corresponds to the minimum critical value is determined. The polarization of the corrected S – wave (δ) is calculated from the following equation:

$$\delta = \tan^{-1}\frac{\lambda_1^y}{\lambda_1^x} \tag{3.4}$$

where λ_1^y and λ_1^x the unit vectors defining the first eigenvector of the covariance matrix corresponding to the selected anisotropy parameters. The φ and δ values are then transformed to the ZNE system. In the last part of this function, the errors of the method are calculated (see Chapter 3.2.4). The final output includes the anisotropy parameters, the critical value and corresponding trial arrays, the initial polarization correction, the errors and the 95% confidence level of the critical value (C_{95}) used in the error estimation.

To compute the errors of the EV method, three additional functions, included in the module, are used. Initially, the corrected (for anisotropy) waveforms in the LQT or ZRT system are acquired. The degrees of freedom (ν) are calculated with the *NDF* function. If that number is lower than the number of parameters (two), the threshold critical value is equal to the minimum and a warning is posted. Otherwise, it is obtained from the *confRegion* function. Finally, the anisotropy parameters, trial and critical value arrays, as well as the threshold critical value are passed to the *getBounds* function to calculate the errors.

eigenvalue.py also contains the function *writeEigenvalueResults*, which is used to write the results of the EV methods in an output file. This contains the measurement details (φ, t_d and corresponding errors. polarization correction, selection window and C_{95}), as well as a table of the calculated critical value (energy or λ_2) for every parameter pair.

3.2.5. clustering.py

This module contains the **clustering** class which consists of all the underlying functions used in the cluster analysis. In order to not lock out the main GUI and be able to provide a visual response of the background process to the user, a second thread is employed. Thus, the main thread runs in parallel with a new one, which handles the clustering process. The **clustering** class emits two signals, to interact with the main thread, to notify whether the

procedure was completed successfully or failed. The analysis described in Chapter 2.4 is implemented with the *run* method of the class (Figure 3.3).

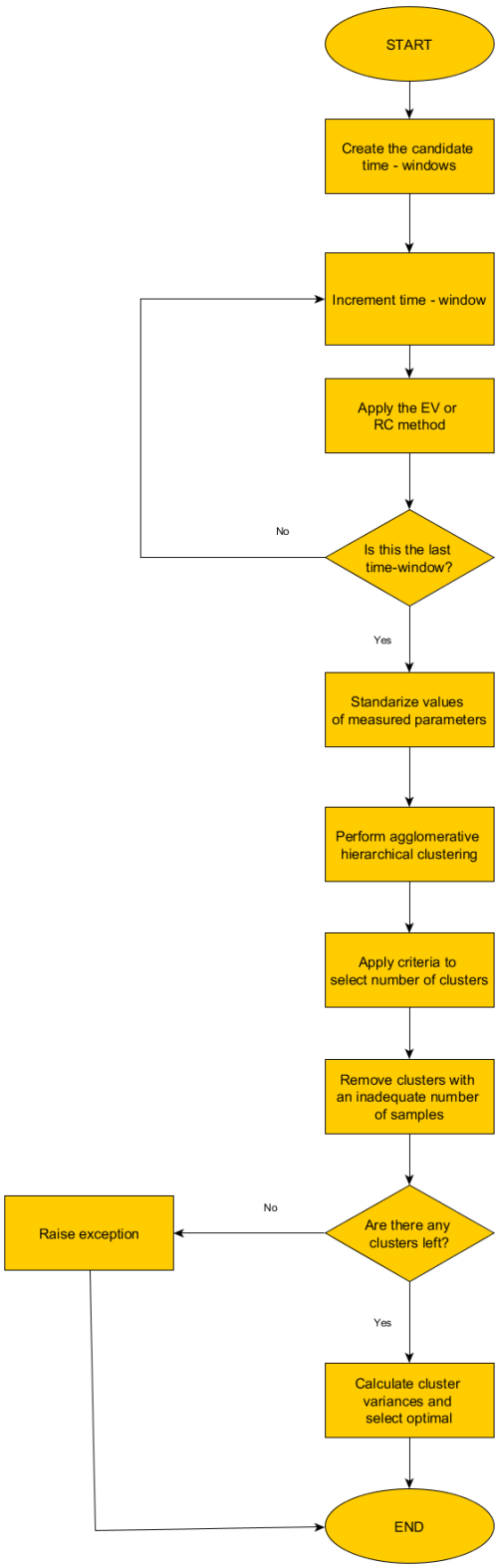

Figure 3.3. Flow chart showcasing the algorithm used in the implementation of the CA method in the Pytheas software.

At first, the windows are defined based on the user provided boundaries. To avoid the use of fixed windows for an entire dataset, a windowing process was adapted from Savage et al. (2010). The user provided parameters are the minimum time intervals before and after the S – arrival, the time steps for the start and end times, the maximum accepted value for the difference between the S- and P- travel -times (S – P), the time window after the S – arrival for the calculation of the dominant period, an extending factor (f_e) to determine the longest accepted window end and the boundaries for the dominant period. The first start point is defined as the half S – P time. The first end point is equal to the dominant period in the S window. The last end point is determined from the dominant period multiplied by the extending factor. To better explain the windowing parameters, these are presented in Table 3.3.

Table 3.3. *Parameters used in the windowing for the CA method. The definition of each is also provided. Where "User" are parameters that are defined in the configuration file. For an explanation of each parameter see Chapter 2.4.*

Parameter	Definition
T_{beg0}	$\frac{t_{SP}}{2}$
T_{beg1}	User
ΔT_{beg}	User
N_{beg}	$\frac{T_{beg1} - T_{beg0}}{\Delta T_{beg}}$
T_{end0}	T_0
T_{end1}	$f_e * T_0$
ΔT_{end}	User
N_{end}	$\frac{T_{end1} - T_{end0}}{\Delta T_{end}}$

The program then iterates over each window and applies the selected SWS analysis method. For each completion, a signal is emitted to the main thread and the progress bar, created before the initialization of **clustering**, is updated, to inform the analyst. Measurements (φ, t_d and their respective uncertainties) are stored in a dictionary that is then converted to four discrete arrays (one for each variable). Values are then normalized, to hold the same weight in the clustering process (given that t_d is in s, the much larger – arithmetically – angle values of φ would dominate). Then, the labels of the data are obtained with the *getClustersSK* method. This function uses the sklearn **AgglomerativeClustering** class to categorize the data in different clusters, for various numbers of clusters M in the range between 1 and the maximum number of data points. The labels (i.e. the cluster identification) for each data point and for each number of clusters are stored in a dictionary. To facilitate the rest of the analysis, observations are stored – per cluster per number of clusters – in a new array, by using *getClusterDataSK*. At this stage, the criteria for selecting M are applied. The Duda and Hart (1973) score (c_{DH}) is first computed from the *critDudaHartSK* method, for M in an ascending order. In other words, starting for $M = 1$, c_{DH} is calculated. For the first M that satisfies the condition $c_{DH} \geq 3.2$, the loop breaks and $M_{max} = M$. If no eligible c_{DH} is found, then M_{max} is equal to the number of observations. Next, the criterion of Caliński and Harabasz (1974) must be applied. For this purpose, the *critCalHara* method is used, which, in turn, utilizes the *metrics.calinski_harabaz_score* from sklearn. A score (c_{CH}) is obtained for every M In the range $[M, M_{max}]$. A user specified

maximum threshold for the number of clusters is used to remove very high values. The M with the highest score is the accepted one. Clusters with a number of observations less than the one set by the analyst are rejected as spurious. Following, the within – cluster and mean data variances are computed. The cluster with the smallest variance is the optimal one. Finally, the anisotropy parameters and their uncertainties, as well as the various quantities used in the creation of the quality control figure (described in the next paragraph) are set as member variables of **clustering**, to permit use in the main thread.

The "clustering" module also contains a function to plot a quality control figure (*clusterQCDiagram*). The user can use this to determine whether the clustering is appropriate and adjust the configuration as needed.

3.2.6. tools.py

The "tools" module is a collection of various functions that do not belong in other modules. *lengthcheck* performs an examination on the length of the timeseries provided. This is essential for any SWS software, as even one sample difference between the two horizontal components can lead to erroneous results. For the two given traces, their start time and end time are synchronized. The correction is performed by adding zeroes in the edges of the signal, to ensure that the minimum start time and maximum end time are used. This is achieved through the *trim* function of Obspy (with a padding value of zero). *timedelay* is used to add the measured t_d to a waveform. For positive time – delays, zeroes (equal to the number of samples that correspond to t_d) are inserted at the start of the signal. If it's negative, a number of samples (equal to t_d) is removed from the beginning of the trace.

getTheorArrivals is a function that calculates theoretical arrival times for a given station – event pair with the use of the TauP implementation in Obspy. A velocity model and a StationXML file must be specified by the user in the configuration file "etc/options/taup.cnf". The velocity model must be in a format readable by the Obspy subroutines (e.g. in the "named discontinuities" structure) or be one of the default models (e.g. IASP91). If the model is not one of the included, the file is converted automatically to the npz format. All arrival times of the various phases for the given station – event pair are then calculated and stored in a dictionary. Finally, the *spectrum* function simply calculates the power spectrum of the defined signals. The periodogram method is used for a hanning window. The spectrum is used by Pytheas to automatically obtain the dominant frequency near the shear – wave arrival and recommend the filter boundaries for analyzing the signal. This module also contains the functions responsible for creating the particle motion diagrams (*polarigram* and *hodogram*), the 2D and 3D rotation matrices creation (*Rmatrix2D* and *Rmatrix3D*), as well as the automatic grading function (*autoGrading*).

3.3. Secondary and Quality of Life Features

Other than the main functionality, the aim of the Pytheas software is to provide convenience features by exploiting the capabilities of the Python language and the Obspy toolkit. As such, there are noteworthy features that do not affect the basic SWS processing, but offer a smoother user experience.

3.3.1. Grading Algorithm

A crucial part of any analysis is the evaluation of results. The user can manually select a grade for the result (independent of the processing method used). The current author suggests the manual assessment scheme of Kaviris et al. (2018a) as a suitable system for assigning quality grades (as described in Chapter 2.1). However, an automatic grading system is also offered, to complement the integration of the (semi- or not) automatic methods. Wüstefeld et al. (2010) proposed an intricate method of grading results by combining the EV and RC methods and comparing the parameters produced from both. This approach is largely dependent on both methods and, consequently, includes any issues that may arise during the implementation of either. The main advantage of their procedure is the accurate identification of null measurements (i.e. situations where shear – waves were not split, either due to an isotropic medium or to the shear – wave polarization being parallel to an anisotropic symmetry axis). Savage et al. (2010) consider measurements with a φ sub – parallel or sub – perpendicular to the isotropic S polarization as null, while implementing a grading system based on the SNR, clustering metrics and the SWS parameters.

For the Pytheas software, a new grading algorithm was developed, to function as an initial filter of unsuitable measurements. The algorithm is designed to work on all methods, manual or otherwise. It is based on the assumption that the measurement quality can be quantified by evaluating three parameters; the SNR and errors of φ and t_d. The user must provide specific limits for the three, i.e. the minimum threshold for the SNR and the maximum permitted bound for the errors. The algorithm firsts checks the difference between the isotropic polarization (δ) and φ to determine whether the measurement is considered null, as seen in Savage et al. (2010), based on a maximum accepted difference $\delta_{diff} = |\delta - \varphi|$. Then, a quality score is calculated from three individual ones. These are:

$$\begin{aligned} s_{SNR} &= \frac{SNR_{min}}{SNR} \\ s_{\varphi} &= \frac{\delta\varphi}{\delta\varphi_{max}} \\ s_{t_d} &= \frac{\delta t_d}{\delta t_{d_{max}}} \end{aligned} \tag{3.5}$$

where s the score for each parameter and δ the errors of φ and t_d. If any of these metrics is greater than 1, the measurement is rejected and assigned a cumulative score of infinity. The final score (s) is:

$$s = s_{SNR} + s_{\varphi} + s_{t_d} \tag{3.6}$$

In essence, the cumulative score showcases the degree of the effect the three parameters have. A score of 1.0 is the equivalent of considering that two out of three parameters are within accepted bounds. A score that tends to 0 is an indication of an excellent measurement, while tending to 3.0 (the maximum attainable score) means that the quality of the measurement is low. A grade is then assigned for different situations:

- Grade "A": $s \leq 1.0$
- Grade "B": $1.0 < s \leq 1.5$
- Grade "C": $1.5 < s \leq 2.0$
- Grade "D": $2.0 < s < 3.0$
- Grade "E": rejected measurements
- Grade "N": null measurements

The limits that bound the grading can be tweaked by the user as seen fit. Even though, as stated above, this algorithm can be used for all SWS methods, it is not really suitable for the PM technique, where the errors of the anisotropy parameters are fixed and, as a result, their contribution to the final score is biased based on the limits specified. This could be partially countered by setting limits equal to the errors (thus, s_φ and s_{t_d} being always 1.0) and refining the grading boundaries to take into account variations of the SNR.

In any case, visually assessing the results is still an important step in the quality evaluation of measurements. Other factors, such as waveform matching after the correction and cycle skipping, cannot be quantified from the above algorithm.

3.3.2. Integration of the QUAKEML and STATIONXML schema

eXtensible Markup Language (XML) files are textual files that contain information in a very specific and structured manner. Every XML file is accompanied by a well – defined syntax. This syntax comprises the XML schema. Different formats are defined by different schemas. The QuakeML schema (Schorlemmer et al., 2011) was first released in 2004 and its main goal is to facilitate data exchange between seismological data providers, who often maintained their own standards, requiring conversions between formats. The main information contained in a QuakeML is the metadata concerning an event, such as the origin time, hypocenter coordinates and magnitude information. Furthermore, it can contain information about discrete picks, arrivals, picking and location methods, from multiple institutions. QuakeML is also extensible, permitting the definition of new fields for information not included in the original (e.g. focal mechanisms). One of the planned updates for Pytheas will offer the ability to output any measurements (including error estimation) from SWS to a QuakeML.

StationXML is a similar concept of an XML schema that focuses on station metadata. These are files that, compared to classic formats, contain richer information about a station. In a StationXML file, other than the base information (geographical coordinates), one can also find a concise station history detailing instrument changes. There are fields that contain information about each individual component and channel, as well as the full instrument response stages. Thus, a StationXML file is layered as follows: (1) Network, (2) Station, (3) Channel and (4) Response. In Pytheas, this file is used to obtain the location of each station. Because of that, the information required is up to the Station layer.

3.3.3. Configuration Files

The AR – AIC picker configuration file (picker.cnf) is used to define several variables required by the Obspy *ar_pick* function. First, the frequency band for the initial Butterworth bandpass filter must be set (properties f1 and f2). This filter is applied before the start of the algorithm. Second, the duration (in seconds) of the STA and LTA windows are specified in the properties lta_p and sta_p. These values are utilized in the STA/LTA filter applied to the data to obtain an initial estimation of the P – wave arrival. Additionally, the STA and LTA window duration for the STA – LTA filter used in the S – wave pick must be set (properties sta_s and lta_s). The coefficients of the auto – regressive model to be fit in the data are also specified in the file (properties m_p for P – waves and m_s for S – waves). Finally, the length of the variance window (in seconds) is provided (l_p and l_s, respectively). These parameters are site independent and can be used for large datasets. The user can utilize this file to better calibrate these properties to their needs.

The clustering.cnf file contains settings about the clustering process. In the WINDOWS section, the required properties, as described in Chapter 3.2.5, for the windows definition are set. When using the default values, only the end point will be adjusted severely. It was observed that these parameters led to an adequate estimation of SWS, while not requiring an extensive time period for the calculations. In the CLUSTERING section, the user can set the threshold value for the Duda and Hart criterion. The one proposed by Teanby et al. (2004) is equal to 3.2. The additional criteria of the maximum number of clusters and minimum number of observations per cluster are also defined here.

taup.cnf contains the path to the velocity model to be used in TauP, as well s the path to the StationXML file used by Pytheas. Furthermore, there is a flag to enable or disable recalculation of the angle of incidence. If that value is False, the angle specified in the provided QuakeML will be used.

The grading.cnf file is used to specify the parameters of the grading algorithm. This includes the maximum accepted angle difference for identifying nulls, the thresholds for SNR and the splitting parameters errors, as well as the upper bound of the score of each grade.

3.3.4. Archiving Methods

The Pytheas software is designed to work with a large database of waveforms. To do that efficiently, especially at the stage where the user needs to empirically identify the optimal parametrization, a archiving procedure is implemented. After reading all the event origin times from the input catalogue, the software performs a search for event codes in the given master data directory. The event codes are stored in a Numpy array. To facilitate the computations, the UNIX timestamp (i.e. the amount of seconds since 1970-01-01T00:00:00 GMT) is used instead of the full date and time. To properly match event code and origin time, each time from the catalogue is subtracted from the code array. The absolute difference that is smaller than the maximum accepted threshold (defined at 5 s) belongs to the origin – code pair that best fit.

To avoid searching anew for the event codes and matching them to the origin times, the results of the above procedure are stored locally. The file “etc/options/paths.exp” contains the paths for the catalogue file and the master data folder last used. “etc/index” contains two types of files. The index files with the “.idx’ extension (and named after their creation date)

include all the paths of event folders acquired from previous runs. These files are matched to specific master data directories in the “index.cat” file. Also, the matching between event codes and origin times is stored in the “/etc/evcor/evcor.dat” file. Thus, all this information is loaded at the startup and the software does not need to execute the event data queries again, saving up considerable time.

3.4. Example of the Automated Process

To showcase the use of the Pytheas software, an event from the WGoC dataset (Chapter 4.1) is exhibited. This is the same example used to showcase the PM method (Chapter 2.1). The SNR is 5.3. After picking the S – arrival, the option "Cluster Analysis (EV)" from the "Splitting" menu is selected. The software then displays a progress bar showing the percentage of windows that have already been analyzed. When the analysis is completed, two figures are displayed, to aid the analyst in evaluating the measurement. One refers to the EV and the other to the CA method.

Concerning the first (Figure 3.4), to the left of the figure, the filtered horizontal components are shown in the following order: (1) the original north and east channels, (2) the fast- and slow- rotated components, after applying the time – delay correction and (3) the corrected for anisotropy north and east waveforms. Below that, the particle motion diagrams exhibit the motion before (solid line) and after (dashed line) the correction. Finally, the contour plot to the right shows the distribution of the normalized (according to the 95% confidence level value) critical value, λ_2. The bold contour line traces the 95% confidence interval, while the selected value is marked by the dashed cross. The parameters pair for this case is N90.0°E for φ and 80.0 ms for t_d, with errors of N1.29°E and 0.59 ms, respectively. The isotropic polarization is N55.8°E. This measurement is automatically graded as "A". This example showcases an angle of incidence outside the shear – wave window, due to its initial selection with the estimate seen in Chapter 4 (which calculated an angle of 34.1°). The shown angle is obtained from TauP.

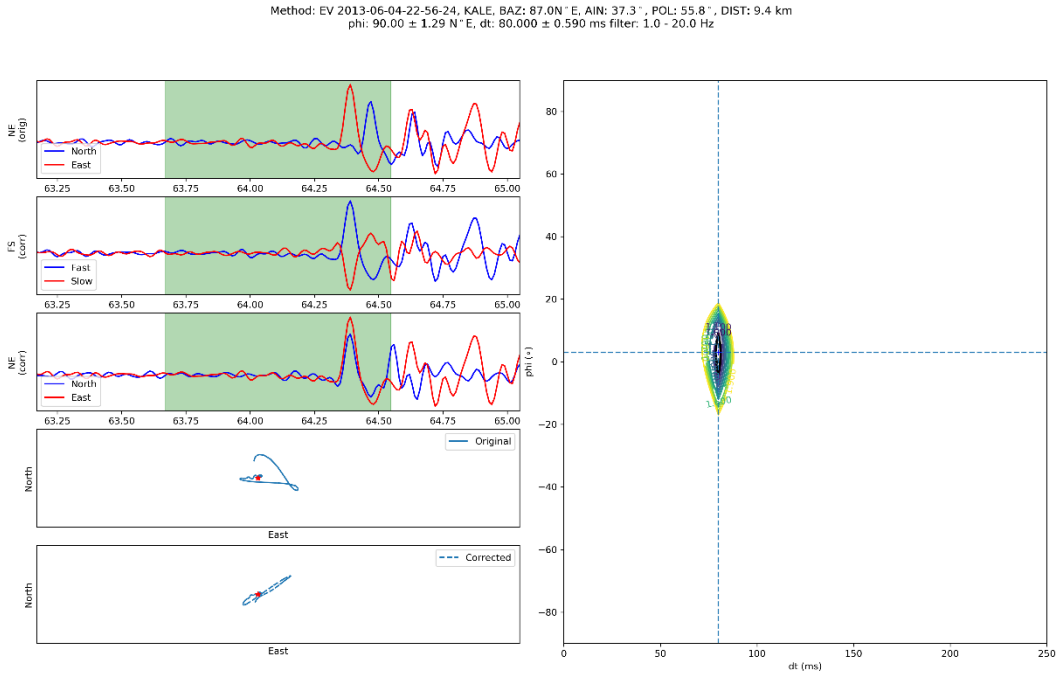

Figure 3.4. Quality control figure to evaluate the results of the EV method. The title includes the following information: method type, origin time, station name, backazimuth (BAZ), angle of incidence (AIN), epicentral distance (DIST), φ, t_d (and their respective errors), isotropic polarization (POL) and the filter band. For this example, the waveforms are filtered in the 1 – 20 Hz range. The unique 95% confidence level contour indicates that there is no cycle skipping in the measurement. The linearization of the particle motion and the matching between the corrected horizontal waveforms are further indication of the excellent quality of the measurement.

The CA quality control figure (Figure 3.5) is used to evaluate the steps of the clustering. At the top left, the initial data (i.e. measurements for each window) are shown. Next (top right), the variation of the Calinski – Harabasz criterion is shown, with the selected number of clusters marked by the crosshair. The final clustering (after finding the number of clusters) is shown at the middle left. The selected measurement window (which belongs to the optimal cluster) is shown at the middle right subplot. The two bottom plots show the stability of the measurement for each window, for both φ (left) and t_d (right). Results are written in the "spl" format file and include the values of the optimal measurement window. In this case, the window ranges from 63.67 to 64.55 s after the start time of the waveform. The solution in this window is shown in Figure 3.4. The parameters obtained from the fully automated process are identical to those determined from the manual analysis (see example in Chapter 2.1).

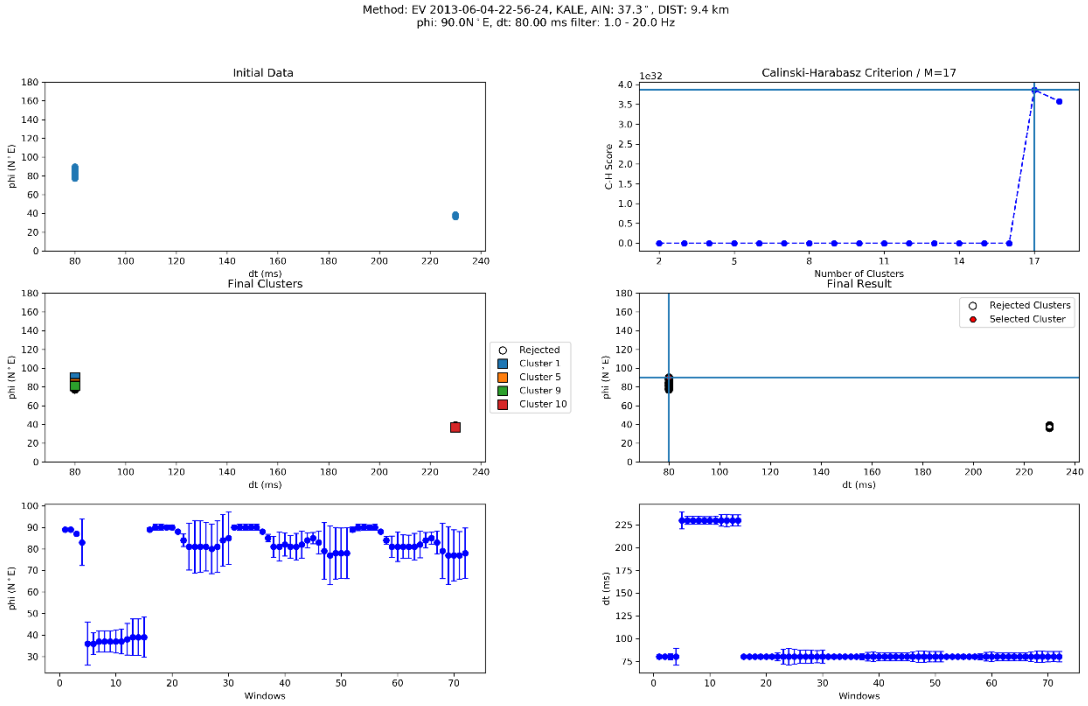

Figure 3.5. Quality control figure to evaluate the results of the CA method. The title includes the following information: method type, origin time, station name, angle of incidence, (AIN) epicentral distance (DIST), φ, t_d and the filter band. For this example, the waveforms are filtered in the 1 – 20 Hz range.

4. Case Studies

To evaluate the performance of the Pytheas software and investigate the viability of the EV and CA methods, waveform data from earthquakes originating in three different regimes were used: (1) the WGoC (tectonic regime), (2) the SVC (volcanic regime) and (3) the WOF (induced seismicity).

A similar procedure scheme was employed for all three cases. For the WGoC and the SVC, the phase arrivals were collected from the Seismological Laboratory of the National and Kapodistrian University of Athens (SL – NKUA) database and the hypocenters were determined with the Hypoinverse (Klein, 2002) program. The velocity models used in the location are presented in Table 4.1. For an initial estimate of the angle of incidence (i_0), the following relation was used (Kapetanidis, 2017):

$$i_0 = \sin^{-1}\left(\frac{V_{top}}{V_{bot}}\sin(i_h)\right) \tag{4.1}$$

where V_{top} and V_{bot} the velocities at the surface and the focus, respectively, and i_h the takeoff angle. Only upwards travelling rays were considered (i.e. $i_h > 90°$). The angles of incidence for each ray during the analysis were calculated with the TauP algorithm through the Pytheas software, using the full velocity model. For the WOF dataset, event and station details were used as provided by the University of Kansas (KU), including incidence angles. After visually inspecting a sample of the events to analyze, it was decided to use a universal 1 – 20 Hz Butterworth bandpass filter for processing. SWS was analyzed with the CA and EV methods, using the fully automated catalogue option. The maximum trial time – delay was set at 250 ms, according to relevant literature (Kaviris et al., 2015, 2017, 2018d; Nolte et al., 2017). The fast axis metrics were calculated with circular statistics (Berens, 2009).

Table 4.1. Velocity models used in the current study for the WGoC (Rigo et al., 1996) and the SVC (Papadimitriou et al., 2015).

WGoC		**SVC**	
Ceiling (km)	**V_p (km/s)**	**Ceiling (km)**	**V_p (km/s)**
0.0	4.8	0.0	4.0
4.0	5.2	1.0	4.3
7.2	5.8	3.0	4.8
8.2	6.1	8.0	5.2
10.4	6.3	12.0	5.5
15.0	6.5	16.0	5.8
30.0	7.0	18.0	6.1
-	-	22.0	6.6
-	-	26.0	7.1
$V_p/V_s = 1.80$		$V_p/V_s = 1.81$	

Only graded "A" and "B" measurements were used. For the WOF data, due to the limited amount of results, "C" grades were included as well. In any case, stations with less than 10 initial results were removed from the final catalogues. The criteria used in the grading were: $SNR_{min} = 2.5$, $\delta\varphi_{max} = 10°$, $\delta t_{d_{max}} = sps$ and $\delta_{diff} = 10°$.

4.1. Western Gulf of Corinth, Greece

4.1.1. Geotectonic Setting

The Gulf of Corinth is one of the most tectonically active areas in Europe (Makropoulos et al., 2012; Kouskouna and Sakkas, 2013) and consists a "natural laboratory" for tectonic and seismological studies. It is a complex asymmetrical graben (Armijo et al., 1996; Brooks and Ferentinos, 1984), whose formation started in the Middle Miocene, mainly affected by the local extensional regime in the Aegean, with an arc – parallel direction, as well as the African – Eurasian plate boundary. During the Upper Miocene - Lower Pliocene, the basin evolved into its current state (Papanikolaou and Royden, 2007). The rifting process probably started to the north, where the Aegean extension resulted in crustal thinning (Tiberi et al., 2000). In its western part (WGoC), fault systems are generally striking E – W, dipping antithetically, southward at the northern and northward at the southern coast (Hatzfeld et al., 2000; Moretti et al., 2003). This system is responsible for uplifting of the coast of North Peloponnesus (Tselentis and Makropoulos, 1986; Skourtsos and Kranis, 2009). The extension is mainly constrained in the inner part of the gulf, with deformation measurements of 16 mm/yr (Avallone et al., 2004) and 14 mm/yr (Briole et al., 2003) in a NNW-SSE direction. Beneath the WGoC, a low – angle detachment zone has been suggested at a depth of about 10 km, with a northward 15° dip (Rigo et al., 1996) which is connected to the overlying faults. Those showcase greater dip values, which tend to be reduced northward (Godano et al., 2014). This structure was proposed to facilitate interpreting the accommodation of strain in the gulf, based on observations of microseismicity. Hatzfeld et al. (2000) argued that microseismicity is located in the transition zone between brittle and ductile deformation in the crust. Latorre et al. (2004) observed a zone of crustal weakness at a similar depth. Nevertheless, other researchers have debated that any slip that occurs on the suggested detachment fault is negligible and, as such, the strain can be attributed to the activity of the overlying steep faults (McNeill et al., 2005; Bell et al., 2008). According to Ambraseys and Jackson (1990), the length of the individual faults does not exceed 20 km, which is the cause of the absence of very strong earthquakes ($M_s > 7.0$) in the area. This segmentation, albeit rendering the tectonic evolution of the gulf complex, does not permit the occurrence of such strong events. In addition, surficial traces extend to a maximum of 10 km (Doutsos and Poulimenos, 1992). The segmentation of faults in the gulf is related to the expression of seismic energy mainly through seismic swarms (Potanina et al., 2011; Karakostas et al., 2012; Duverger et al., 2015; Kapetanidis et al., 2015; Mesimeri et al., 2016, 2018). Such episodes were recorded during 2013 and 2014 (Chouliaras et al., 2015; Kapetanidis et al., 2015; Kapetanidis, 2017; Kaviris et al., 2017, 2018a, 2018d), where over 3,500 earthquakes were identified (Figure 4.1). There are several fault systems in the northern and southern coast of the WGoC. Following, the most important ones are described.

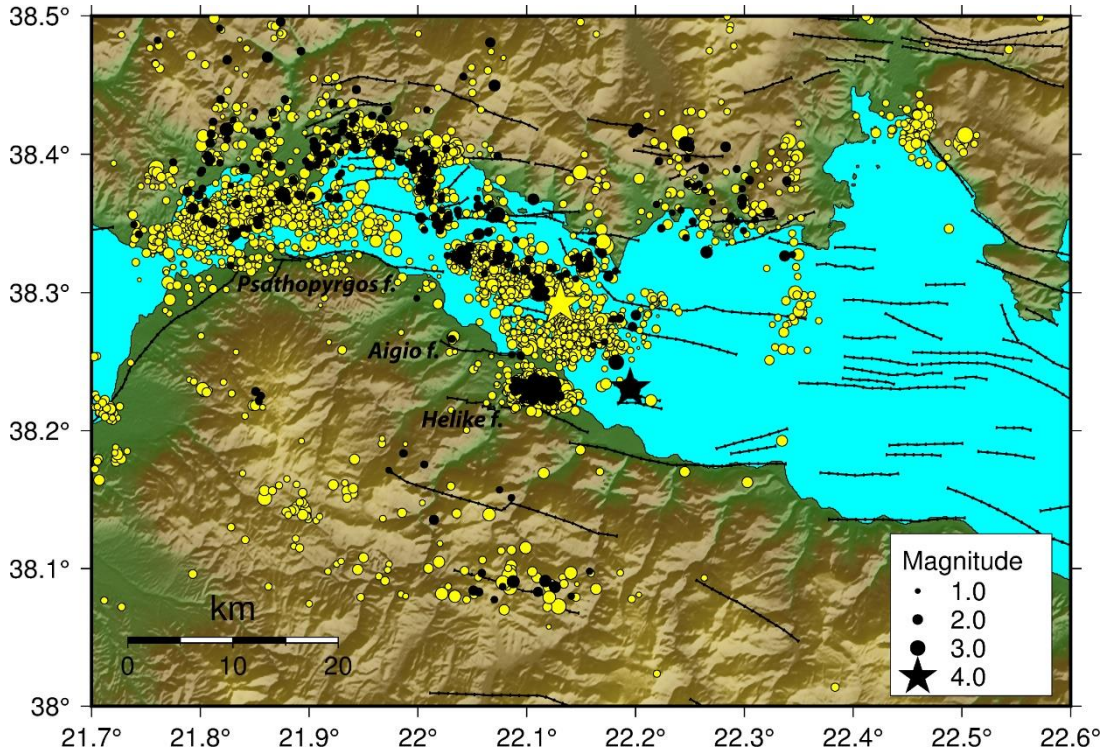

Figure 4.1. Seismotectonic map of the WGoC. The earthquakes that occurred during 2013 and 2014 (yellow) and the ones that provided results during the SWS analysis in at least one station (black) are shown. The faults (black lines) are after Ganas et al. (2013a). For all maps, the topography was obtained from the Shuttle Radar Topography Mission (Reuter et al., 2007). Maps are plotted with the General Mapping Tools (Wessel and Smith, 1991) software.

The Psathopyrgos fault is located at the westernmost part of the area (Figure 4.1), on the southern coast. The broader region of the fault exhibits high mean topography, a strong indication of intense tectonic activity (Tsimi et al., 2007). Dating of corals in marine deposits led to the estimation of an uplift rate, for the southern block, equal to 0.7 mm/year (Houghton, 2003). The Aigio fault (Figure 4.1) is located near the homonymous city, east of the Psathopyrgos fault. It remains active since the historical era (Pantosti et al., 2004) and has presented significant seismicity in recent years (Mesimeri et al., 2016). It extends several kilometers to the east and continues offshore, across the shelf (McNeill et al., 2005). While it is severely segmented, a mean strike of N100°E can be deduced (Koukouvelas and Doutsos, 1996), as well as a length of 12 km (Koukouvelas, 1998). Evidence from marine deposits suggest a slip rate between 9 and 11 mm/year and an uplift rate in the range of 1.05 – 1.20 mm/yr (De Martini et al., 2004). The Aigio fault was the focus of multiple studies due to the occurrence of the 15th June 1995 (M_s = 6.2) earthquake, in the vicinity. Even though this event has been attributed to the Aigio fault (Koukouvelas and Doutsos, 1996; Koukouvelas, 1998), Bernard et al. (1997) suggested that a low – angle (uncommon for the area) fault was the causative one. Finally, the Helike fault (Figure 4.1), to the SE of Aigio, is one of the most important structures in the gulf. Its activity has been identified since the early Copper Age, accompanied by destructive earthquakes, as the famous 373 B.C. event which led to the generation of a tsunami and the devastation of the ancient city of Helike (Soter and Katsonopoulou, 2011). The slip rate of the northern hanging wall is not constant and it has been proposed that slip on the fault is conducted with variations of high (2.0 mm/yr) and low (0.3 mm/yr) rate. The transition between slip rate periods is denoted by seismic outbursts (Koukouvelas et al., 2005).

Nevertheless, De Martini et al. (2004) deduced significantly higher slip rates which vary spatially, i.e. 7 – 9 mm/yr for the eastern and 9 – 11 mm/yr for the western part of the structure. In recent years, seismicity in the Helike fault has been expressed mainly through seismic swarms (Kapetanidis et al., 2015). At the northern coast of the WGoC, seismicity and tectonic activity are much more moderate. On January 2010, a doublet of events occurred near the city of Efpalio, both with a magnitude of $M_w = 5.1$ (Sokos et al., 2012; Ganas et al., 2013b).

The WGoC is covered by two seismological networks. The Hellenic Unified Seismological Network (HUSN) consists of stations located in the entirety of Greece, installed and operated by the following Greek institutes: (i) SL – NKUA, (ii) Seismological Laboratory of the University of Patras, (iii) Department of Geophysics of the Aristotle University of Thessaloniki, (iv) Technological Educational Institute of Crete and (v) Institute of Geodynamics of the National Observatory of Athens. In addition, a French – Greek initiative in 2000 initiated the Corinth Rift Laboratory Network (CRLN), aiming to study geotectonic processes around Aigio (Lyon-Caen et al., 2004). All stations located in the WGoC are equipped with broadband seismometers, operating at a sampling rate of 100 sps. The CRLN also includes instruments located in boreholes.

Shear – wave splitting has been extensively studied in the gulf. The majority of measurements and research time have discovered an anisotropic layer with a dominant anisotropy direction (WNW – ESE) perpendicular to the rift's extension, oriented according to $\sigma_{H\,max}$ (Bernard et al., 1997; Bouin et al., 1996; Giannopoulos et al., 2015; Kaviris et al., 2017, 2018a, 2018d). This is consistent with an anisotropic medium pervaded by fluid – saturated microcracks. Thus, shear – wave splitting in the WGoC has been interpreted with the APE model. However, deviations from these have been observed in three stations (SERG, PYRG and MALA) at the NW part of the area (Giannopoulos et al., 2015; Kaviris et al., 2017, 2018a, 2018d). An interesting observation is the shift of the S_{fast} polarization direction in PSAR. Measurements from the early 90s documented a direction of ~N64°E (Bouin et al., 1996). However, data from the 2013 and 2014 have shown a very well constrained strike of approximately N90°E (Kaviris et al., 2018d).

4.1.2. Results

For the application of the automatic CA – EV technique, 3,808 events (with 38,130 corresponding S – wave arrivals) in the WGoC were initially selected, during 2013 and 2014, with local magnitudes M_L ranging from 0.3 to 5.1. Further selection criteria excluded any events that did not have at least one arrival with an angle of incidence less than or equal to 35.0°. Due to limitations in resources, only the first semester of 2013 was analyzed. The gargantuan task of automatically processing the full dataset (over 15,510 eligible pairs) is reserved for the future, when more appropriate circumstances arise. Thus, the final catalogue consisted of 1,136 events (18,556 arrivals), with 7,844 rays within the shear – wave window (Figure 4.2) and magnitudes (M_L) between 1.2 and 4.2. It is noted that the selection includes several arrivals in stations that exhibited technical problems and had to be excluded from the analysis. For example, an issue with the North component of the AGEO station enabled analysts to determine the arrival of the shear – wave (in the East channel), but could not be used in the SWS processing. Considering that AGEO is located close to the Helike swarm of 2013, this led to a *de facto* exclusion of a high number of potential measurements.

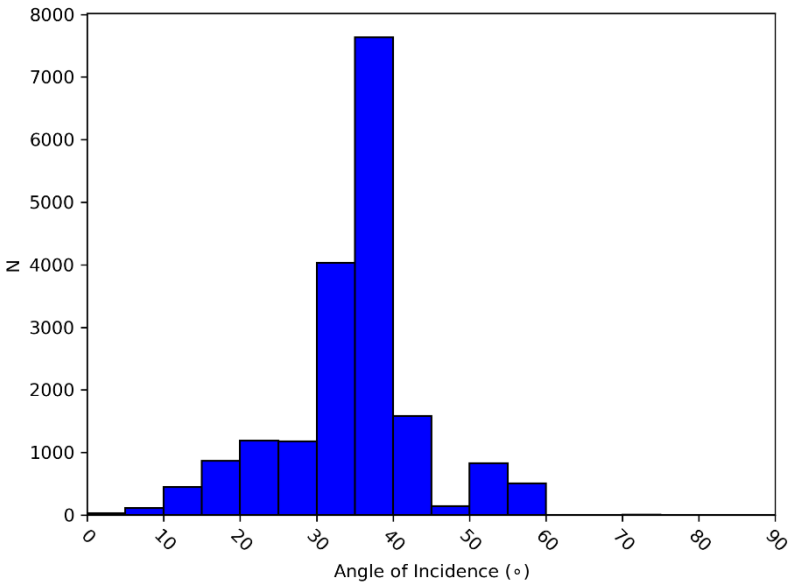

Figure 4.2. Distribution of the angle of incidence in the WGoC catalogue, for S- arrivals. Bin spacing is 5°.

The number of results is 933 parameter pairs out of 375 events. The processing took 16 hours and 36 minutes to complete. Grading led to a final dataset of 358 pairs (grades "A" and "B"). The results for the thirteen stations in the WGoC are presented in Table 4.2. Polarization directions were corrected for the instrument's orientation according to Kapetanidis (2017). Due to the large scatter in the observations, the median value was preferred over the mean for φ. These are shown for each station in Figure 4.3.

Table 4.2. Summary of statistics for analyzed SWS parameters in the WGoC, with the CA and EV methods. In the elevation column, the depth of boreholes is noted within the parentheses. N is the number of observations, $\widetilde{\varphi}$ the median value of the S_{fast} polarization direction, $\delta\varphi$ the standard error of mean of φ, $\overline{t_d}$ the mean time – delay and δt_d the standard error of mean of t_d.

Station	Elevation (m)	N	$\widetilde{\varphi}(N°E)$	$\delta\varphi(N°E)$	$\overline{t_d}(ms)$	$\delta t_d(ms)$
AIOA	198 (130)	66	79	4	48	3
ALIK	37 (70)	29	114	5	102	9
EFP	107	49	70	7	65	7
KALE	760	18	142	15	69	11
LAKA	505	15	123	11	74	11
MALA	7 (180)	21	61	11	59	10
MG00	143	8	65	16	68	17
PANR	149	12	163	11	83	19
PSAR	117	35	103	7	73	8
PYRG	596	27	91	10	81	11
ROD3	448	2	106	16	80	60
SERG	480	41	49	5	53	7
TRIZ	59	35	105	11	63	7

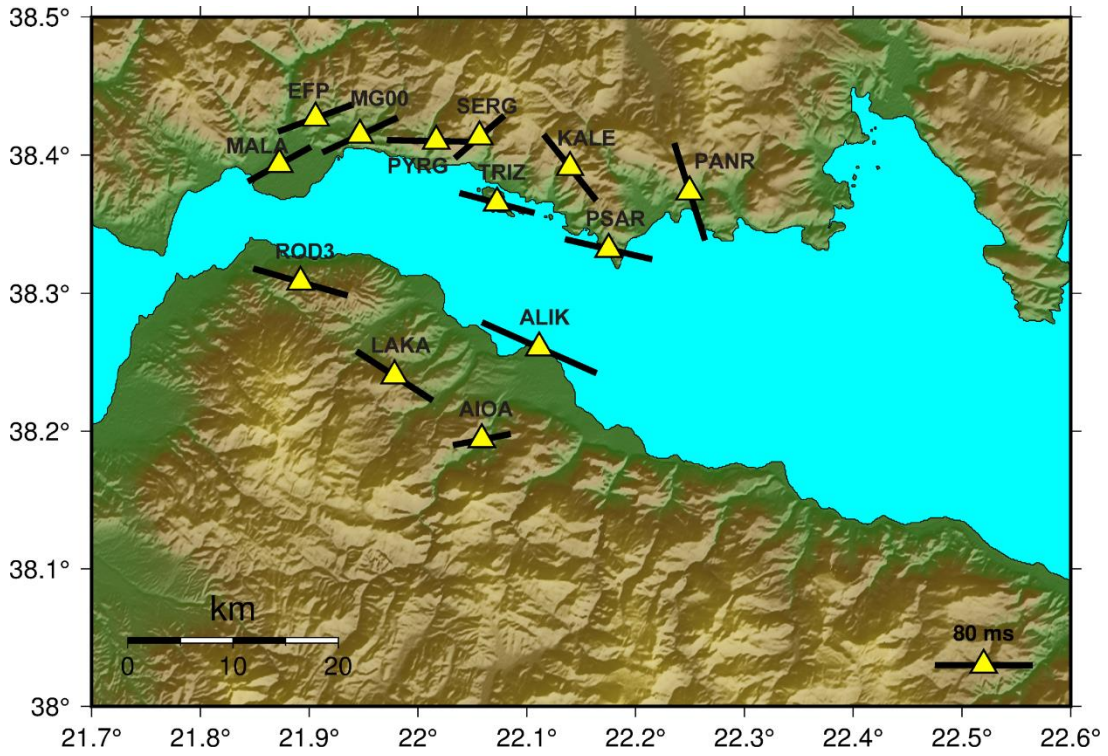

Figure 4.3. Map of the WGoC showcasing the median polarization direction for each station. The bar is proportional to the mean time – delay.

As mentioned, above, there is a significant scatter in the measurements of φ in most stations. The rose diagrams (Figure 4.4) showcase this. Stations EFP, KALE and TRIZ are the most characteristic, with measurements of the polarization equally populating most directions. However, stations AIOA and ALIK showcase surprisingly constrained measurements, reflected in the small errors. Most median polarization values are within 45° of the E – W direction. Deviations from this orientation are observed at stations (KALE, SERG and PANR) that are not located in close proximity to each other.

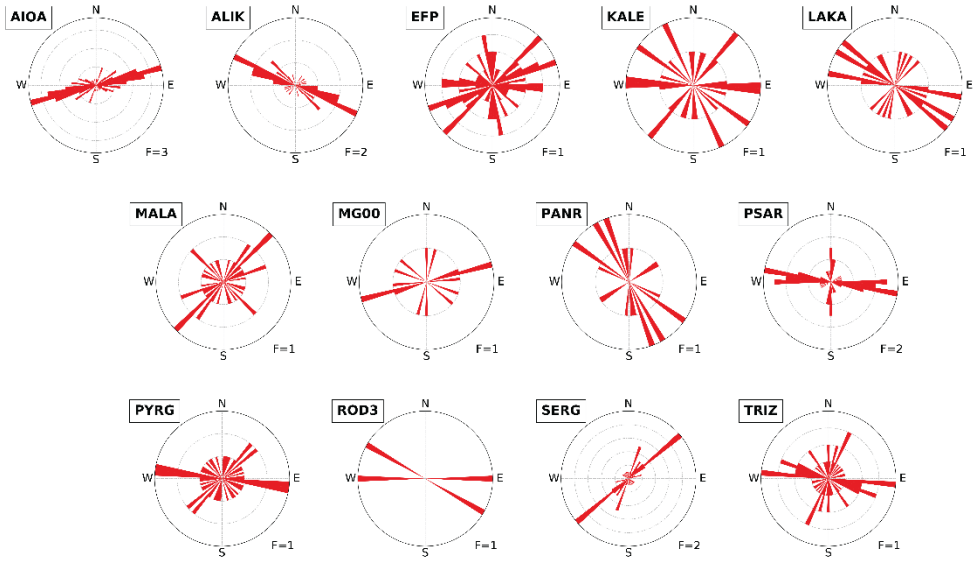

Figure 4.4. Rose diagrams of φ in the WGoC, with the CA and EV methods. F is the number of observations per grid line.

Concerning time – delays, two sub – groups can be identified: below 70 ms and above that. There is no clear indication of any relationship between those stations. Elevation, which theoretically would affect t_d by increasing the ray path and, thus, the time – delay, does not seem to be a factor. In terms of individual values, time – delays are present in the whole 0 – 250 ms range, as expected for a six months period with highly volatile seismicity. Nevertheless, they seem to be concentrated between 30 and 50 ms (**Error! Reference source not found.**).

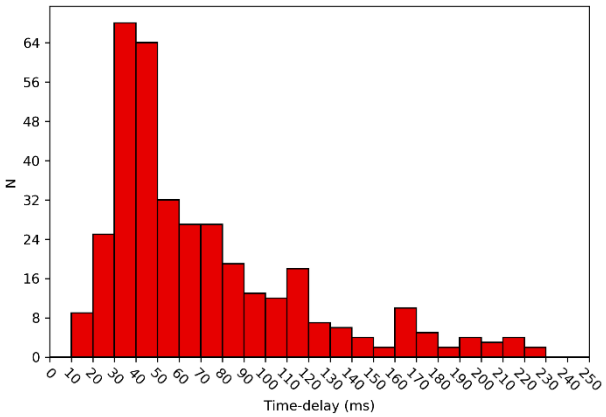

Figure 4.5. Summary time – delay histogram in the WGoC

For the purpose of applying the SWST method of Johnson et al. (2011), through the TESSA software, the SWS parameters catalogue obtained from the current study is not suitable because of the significant scatter observed to both φ and t_d. Thus, the catalogue of (Kaviris et al., 2018d), containing events from both 2013 and 2014, was utilized instead. These results were obtained with the manual PM method. A total of 20 stations and 1,447 events is included, resulting to 2,271 event – station pairs. To obtain the grid for the inversion, the quadtree gridding process was preferred, over a regular grid. Coverage of rays is highly uneven in the area (Figure 4.6) and, thus, a standard grid with a fixed interval of nodes would include a significant number of cells with no data. The minimum size for each grid cell during the quadtree processing was set to 1 km. Only cells with a number of rays between 5 and 50 were accepted. The maximum number of rays is used in the gridding, to split cells with a large number of data.

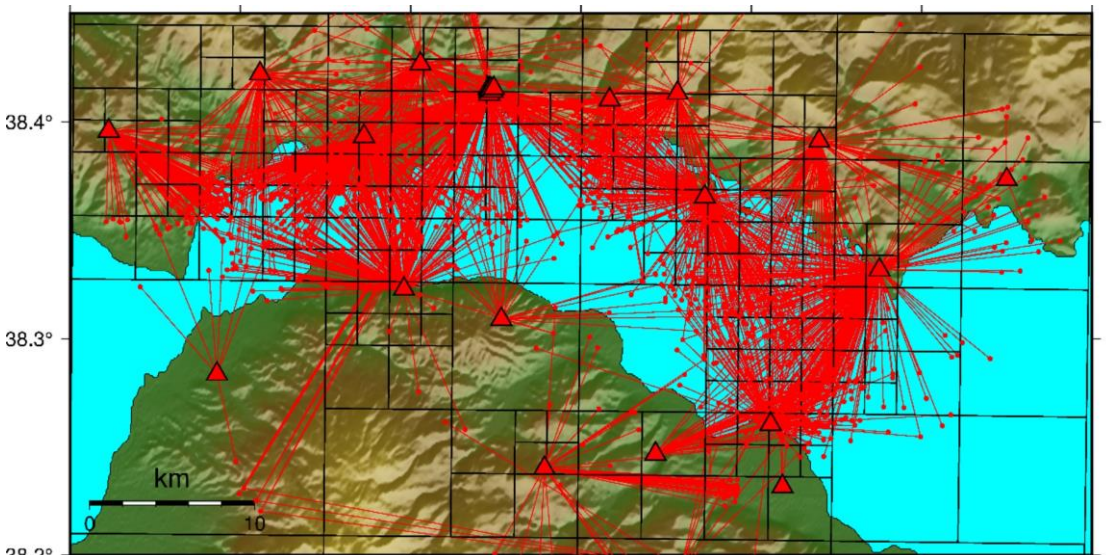

Figure 4.6. The area for which the SWST was applied in the WGoC. The rays (red lines) are connecting events (red circles) and stations (red triangles). The final grid used after the quadtree gridding process is also shown.

To determine areas where the results are constrained, a checkerboard test was applied (Figure 4.7). The size of each checker was set to 1 km. The test revealed an area of reliable results at the center of the study area, as seen in Figure 4.7. The SW, SE and NE regions could not reproduce the checkerboard.

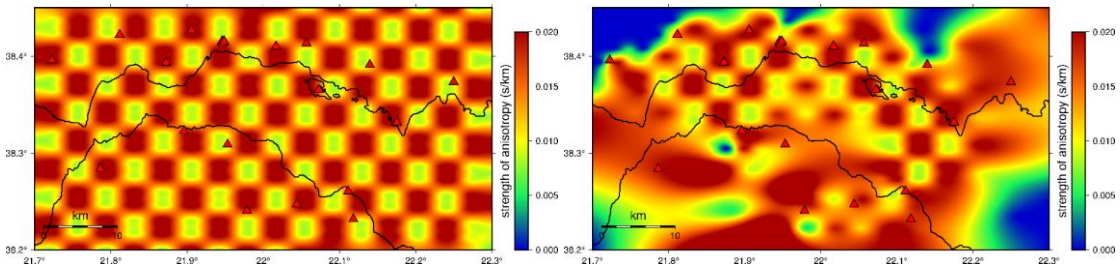

Figure 4.7. Input checkerboard model (left) and synthetic data output (right). Areas where the initial model is recreated are considered to provide adequate resolution. The stations used in the tomography (red triangles) are also presented.

Regarding the tomography, alternating areas of localized high and low anisotropy are observed (Figure 4.8). Most of the central gulf seems to offer values of 10.0 ms/km. The

maximum value observed is 20.0 ms/km. A locus of intense anisotropy is located at TRIZ station. The distribution of the strength of anisotropy does not seem to follow any tectonic features. Considering that shear – wave splitting is attributed to fluid processes in the WGoC (Kaviris et al., 2018b, 2017), this is not surprising. In areas with no resolution, very high values of anisotropy are observed beneath stations. This is probably an artifact, due to the assumption of the method that time – delay is accumulated along the ray path.

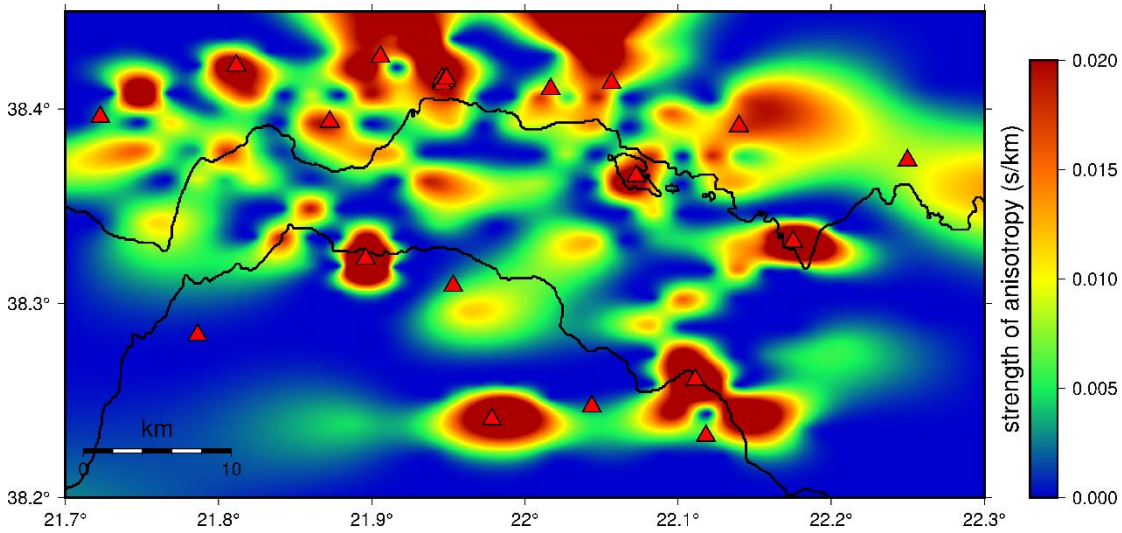

Figure 4.8. 2 – D representation of the strength of anisotropy in the WGoC, produced from the tomographic inversion.

4.2. Santorini Volcanic Complex, Greece

4.2.1. Geotectonic Setting

The SVC is located in the southern Aegean and is one of the most significant volcanic centers of the Greek volcanic arc. The complex is comprised by five islands (Figure 4.9), i.e. Santorini (or Thera), Therasia, Aspronisi, Palea and Nea Kameni, that form a distinct circular shape. The Santorini caldera can be further distinguished into four basins (with a depth range between 290 and 390 m). Activity that belongs to the SVC also occurs at the Columbo underwater crater, which is located to the NE of Santorini at a depth of 500 m (Perissoratis, 1995). In the center of the caldera, the Palea and Nea Kameni islands are the main points of modern volcanic activity in the SVC, with their formation having started around 1200 B.C. (Nomikou et al., 2014). The emergence of the Palea Kameni island was first observed along with the destructive earthquake of 198 B.C. (Papazachos and Papazachou, 2003). The SVC is generally considered a region of low to moderate seismicity (Papadimitriou et al., 2015), as earthquakes are commonly linked to violent volcanic events. Nevertheless, a NE – SW trending zone has been identified between the Santorini and Amorgos islands (which includes Columbo). Seismicity in this zone is much higher and, considering the locality of the magmatic processes, is associated with the development of pathways that facilitate the magma's ascension to the surface (Bohnhoff et al., 2006). The Santorini – Amorgos zone is segmented by transverse NE – SW striking faults (with varying slip) that clearly define three individual secondary basins (Nomikou et al., 2017).

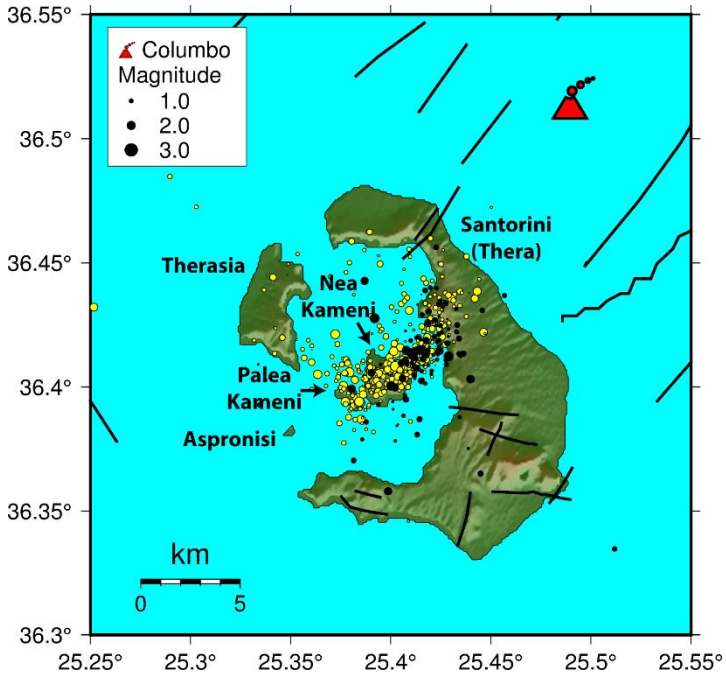

Figure 4.9. Seismotectonic map of the SVC. Shown are the earthquakes that occurred during the unrest period between 2011 and 2012, within the Caldera, (yellow) and the ones that provided results during the SWS analysis in at least one station (black). The faults (black lines) are after Ganas et al. (2013a).

Concerning the eruption history of the SVC, twelve major eruptions have occurred in the past 300,000 years, interrupted by smaller scale ash – producing events (Vespa et al., 2006). The most significant major eruption took place in the 17th century B.C., during the Bronze Age, and is considered a catalyst in the demise of the Minoan civilization (Marinatos, 1939). Simulations have revealed the possible generation of a tsunami wave that could have affected the northern shore of Crete, as well as nearby islands in the southern Aegean (Pareschi et al., 2006). This eruption consisted of five stages and led to the formation of the modern caldera (Druitt, 2014). Tephra was dispersed as far as the Black Sea to the north (Guichard et al., 1993) and Egypt to the south (Stanley and Sheng, 1986). Nevertheless, its importance, both in environmental and historical terms, has been debated (Pyle, 1997).

Between September 2011 and March 2012, seismic activity in the SVC increased dramatically (Vallianatos et al., 2013), suggesting a potential eruption. A significant uplift was observed with a value equal to 14 cm within a year (Foumelis et al., 2013; Lagios et al., 2013). Seismicity was first constrained in the caldera, between the Kameni islands and Santorini. In addition, two earthquakes with magnitudes M_w = 5.1 and M_w = 5.0 occurred to the SW of Santorini in January 2012. Seismological evidence did not suggest an impending volcanic eruption (Papadimitriou et al., 2015). The complex is covered by stations belonging to HUSN.

Shear – wave splitting during the unrest period of 2011 – 2012 was studied by Konstantinou et al. (2013) and Kaviris et al. (2015). Results of the polarization direction were generally consistent, showing mean values between NW – SE and NE – SW. Both research teams attributed the phenomenon to the APE model. Kaviris et al. (2015) featured a more extensive dataset. In addition, even though it was suggested by other methods (e.g. inSAR) that an eruption was within the realm of possibilities, no 90° - flips or equivalent variations in t_n were observed. Indeed, no eruption accompanied the unrest period.

4.2.2. Results

For the EV – CA methods, a catalogue of 520 events (1,520 S - arrivals), located within the caldera, was selected. The dataset spans the unrest period of 2011 – 2012 and the included magnitudes (M_L) range between 0.2 and 3.3. Only 263 rays complied with the shear – wave window criterion ($i_0 \leq 35°$), with the majority being outside (Figure 4.10).

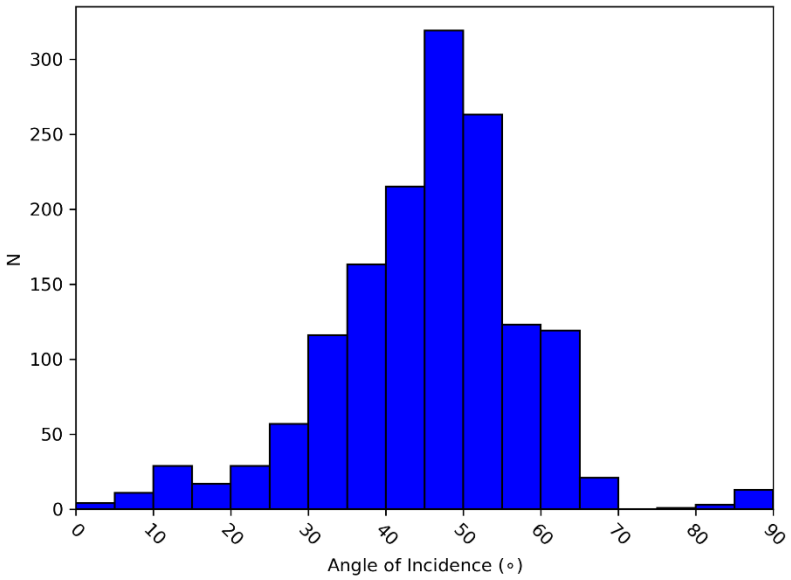

Figure 4.10. Distribution of the angle of incidence in the SVC catalogue. Bin spacing is 5°.

A total of 145 parameter pairs was determined out of 118 events, after 2 hours and 17 minutes of processing. The strict grading criteria rejected approximately half the measurements, with only 82 of them being in the "A" and "B" range. The results for the three stations in the SVC are presented in Table 4.3. The limited number of stations is attributed to the station – event geometry, with the caldera cluster being located closer to the middle part of western Santorini (Figure 4.9). The location of the stations (and their respective anisotropy results) is presented in Figure 4.12.

Table 4.3. Summary of statistics for analyzed SWS parameters in the SVC, with the CA and EV methods. N is the number of observations, $\widetilde{\boldsymbol{\varphi}}$ the median value of the S_{fast} polarization direction, $\boldsymbol{\delta\varphi}$ the standard error of mean of φ, $\overline{\boldsymbol{t_d}}$ the mean time – delay and $\boldsymbol{\delta t_d}$ the standard error of mean of t_d.

Station	Elevation (m)	N	$\widetilde{\boldsymbol{\varphi}}(\boldsymbol{N°E})$	$\boldsymbol{\delta\varphi(N°E)}$	$\overline{\boldsymbol{t_d}}(\boldsymbol{ms})$	$\boldsymbol{\delta t_d(ms)}$
SANT	591	1	114	-	60	-
SNT3	136	15	146	15	83	16
THT1	0	63	171	4	54	6

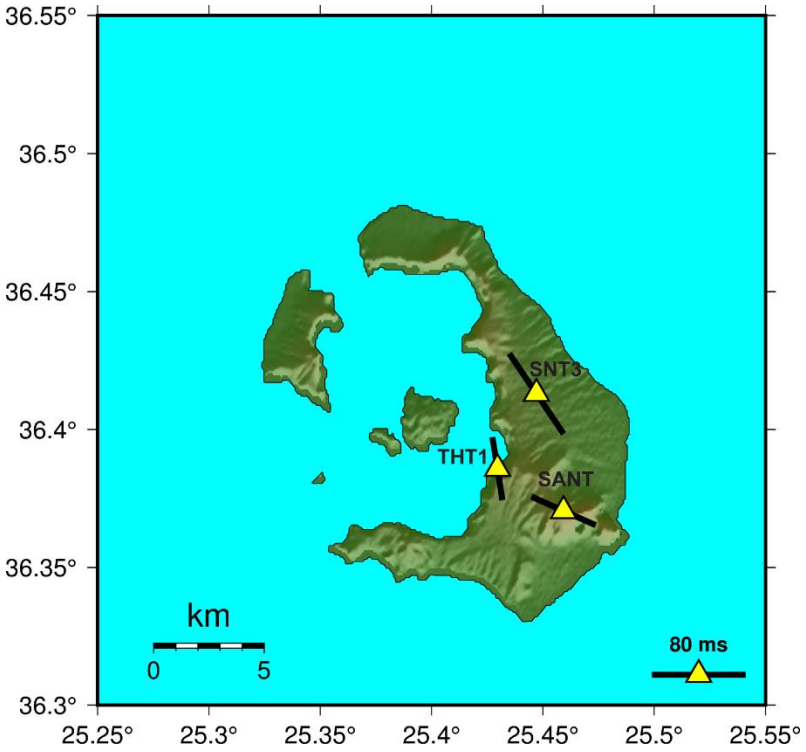

Figure 4.11. Map of the SVC showcasing the median polarization direction for each station. The bar is proportional to the mean time – delay.

The scatter is present yet again (Figure 4.12), but results are generally more constrained compared to some extreme cases in the WGoC. There is only one observation for SANT station, so commenting on it is not reliable. However, both SNT3 and THT1 showcase sub – parallel directions, albeit there is stronger scattering in the former.

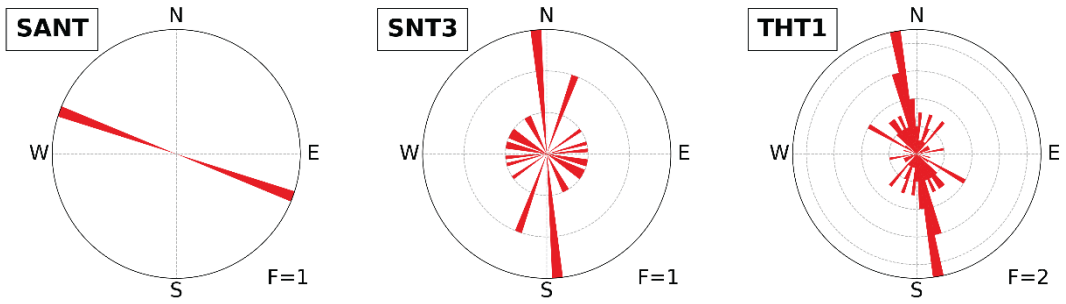

Figure 4.12. Rose diagrams of φ in the SVC, with the CA and EV methods. F is the number of observations per grid line.

Concerning time – delays, SNT3 and THT1 present mean values of a ~30 ms difference. Although it could be argued that the station at the higher elevation features a higher t_d, the difference is rather small (136 m), so this is considered coincidental. A more interesting observation is that SNT3 has both higher mean time – delay and errors. Time – delays are generally constrained within 10 and 90 ms, with some exceptions (Figure 4.13).

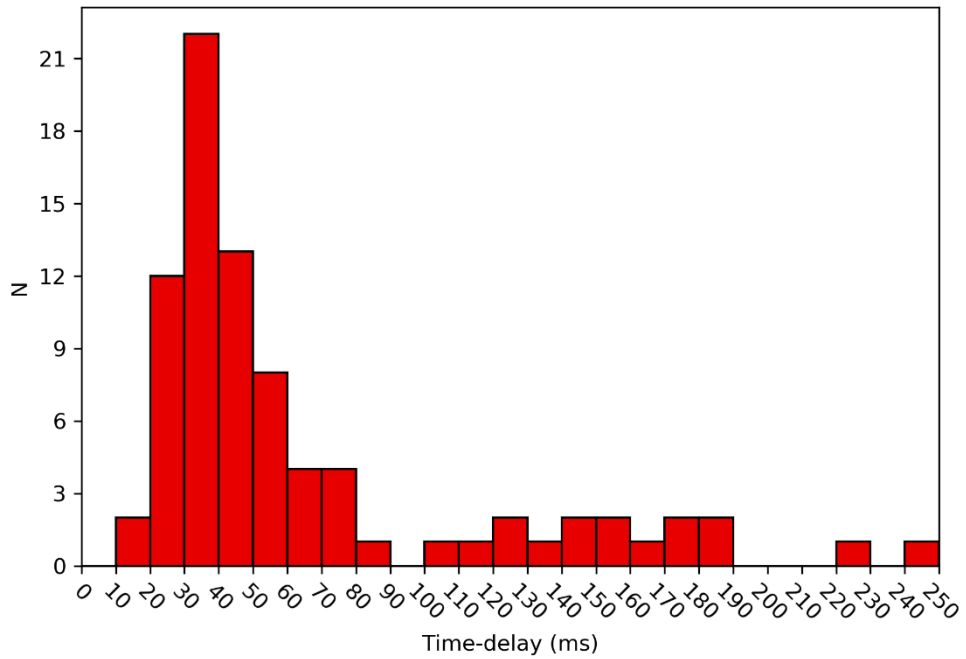

Figure 4.13. Summary time – delay histogram in the SVC, with the CA and EV methods

The SWST for the SVC was employed using the results of Kaviris et al. (2015), where the PM method was utilized. The final catalogue included 6 stations and 217 events, resulting in 355 event – station pairs. As in the case of the WGoC, the quadtree gridding was employed. A minimum accepted cell size of 0.5 km was set, for a number of rays between 5 and 50.

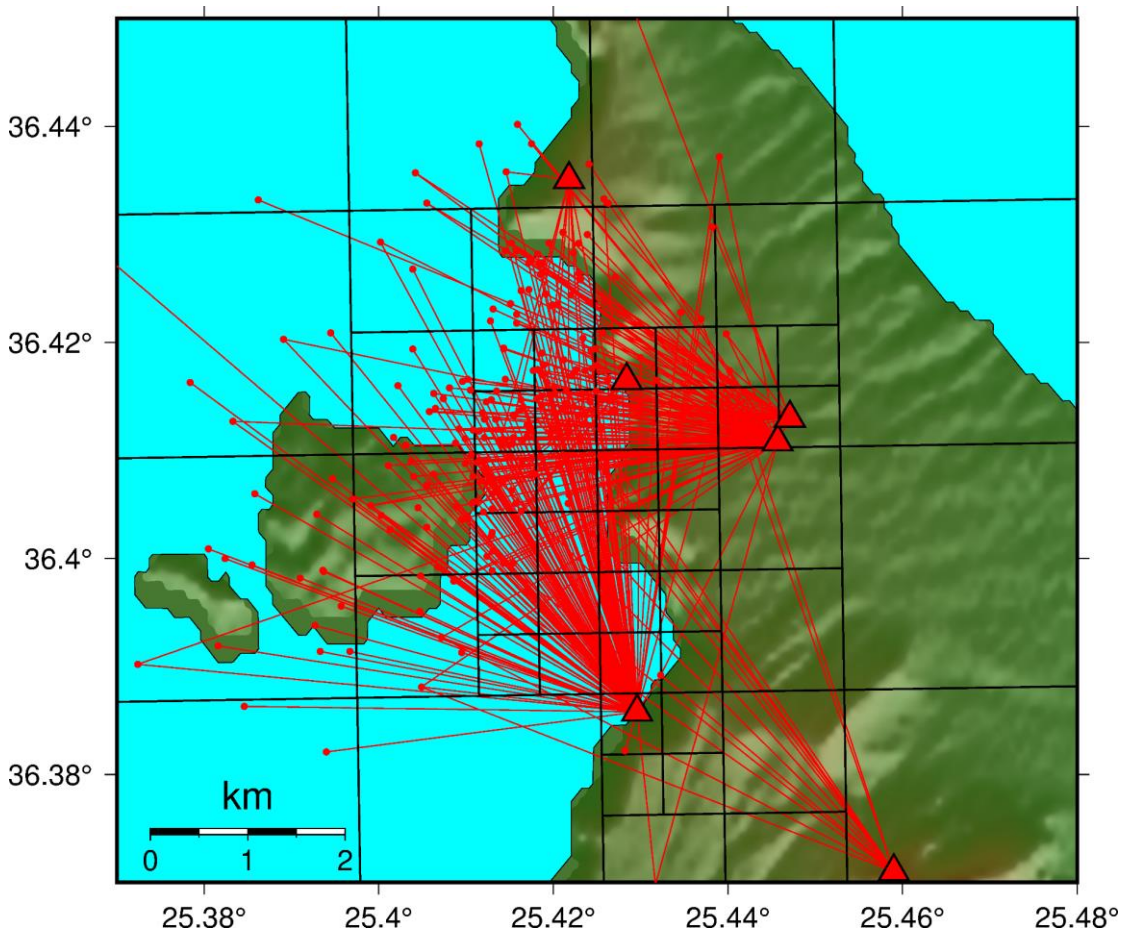

Figure 4.14. The area for which the SWST was applied in the SVC. The rays (red lines) are connecting events (red circles) and stations (red triangles). The final grid used after the quadtree gridding process is also shown.

For the checkerboard tests (Figure 4.15), a 0.5 km interval was selected. While not as clear as the case of the WGoC, resolution is constrained between Nea Kameni and Santorini, where the rays are denser, in an area with NW – SE direction.

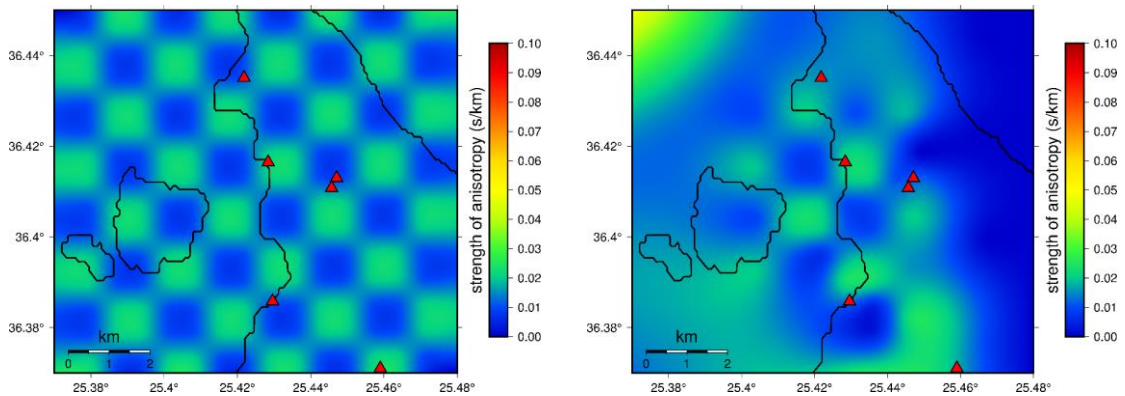

Figure 4.15. Input checkerboard model (left) and synthetic data output (right). Areas where the initial model is recreated are considered to provide adequate resolution. The stations used in the tomography (red triangles) are also presented.

The tomography (Figure 4.16) is characterized by mostly values lower than 60.0 ms/km. The exception is an area to the south of THT1, which anisotropy seems to be much stronger (up 100.0 ms/km). However, this region is arguably within the limits of the resolution and its representation is not considered reliable.

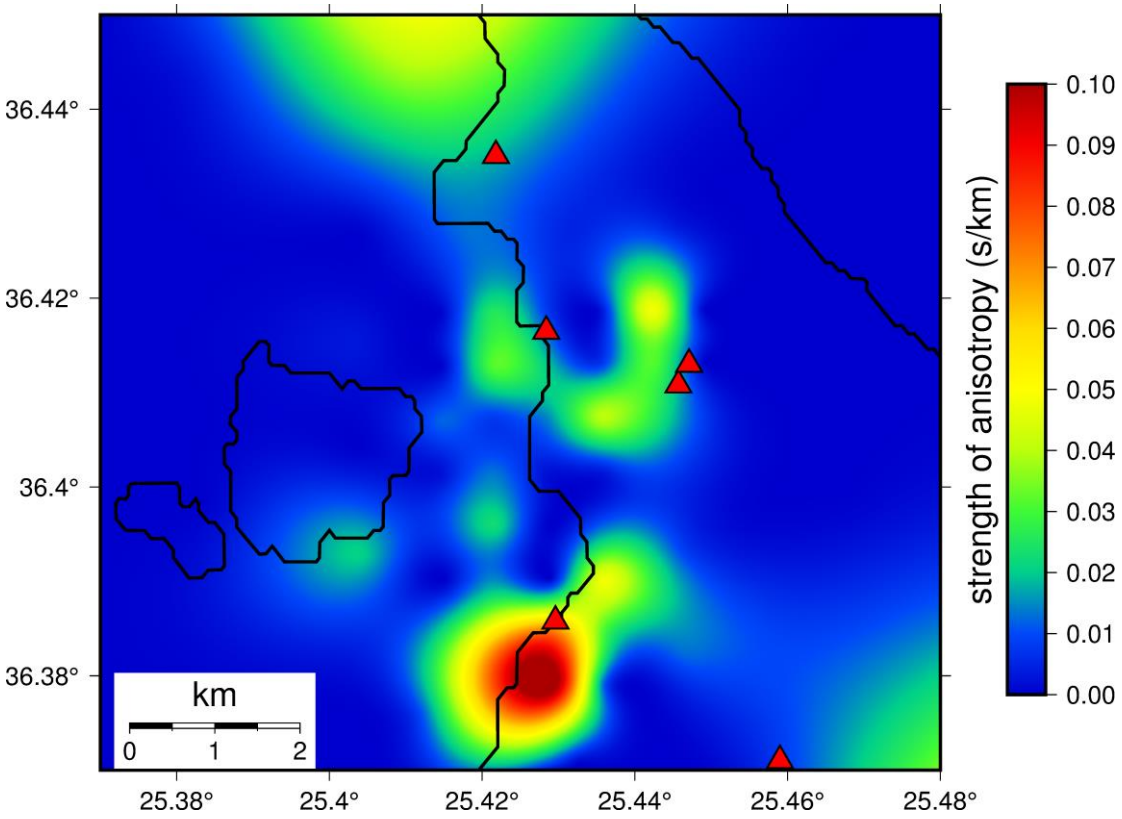

Figure 4.16. 2 – D representation of the strength of anisotropy in the SVC, produced from the tomographic inversion.

4.3. Wellington Oil Field, Kansas, USA

4.3.1. Geotectonic Setting

The WOF is located in the Kansas state in the central USA. The broader area has experienced some seismic activity in the past. Since 2008 there has been an unprecedent increase in both the occurrence frequency and the magnitudes observed. In 2011, a M = 5.6 earthquake occurred in Oklahoma, resulting in moderate damages and injuries. Even weaker events (in the range of 3.0 to 3.9) can lead to minor damage, considering the building code applied in the area (Ellsworth et al., 2015). This surge in seismicity has been linked to oil production – related activity. Wastewater produced during the oil extraction is disposed of by being stored in porous formations bounded by impermeable rocks or faults, which act as seals and block the fluid diffusion. This results in increased pressure in the reservoir which, in turn, can exceed the critical pressure threshold on the faults and trigger slip (Keranen et al., 2013).

Nolte et al. (2017) performed a SWS study in the area of the WOF. They identified two indicators of the wastewater injection affecting the propagation medium. First, a general agreement of φ values with the local $\sigma_{H_{max}}$ (N75°E) was observed between 2010 and 2015. A shift from that orientation was observed in late 2015 and 2016, with polarization being flipped to a general NW – SE direction. Second, time – delays exhibit an increase in variance after the beginning of the injections. In general, the above are strong indicators of a rock volume where pore pressure is increased.

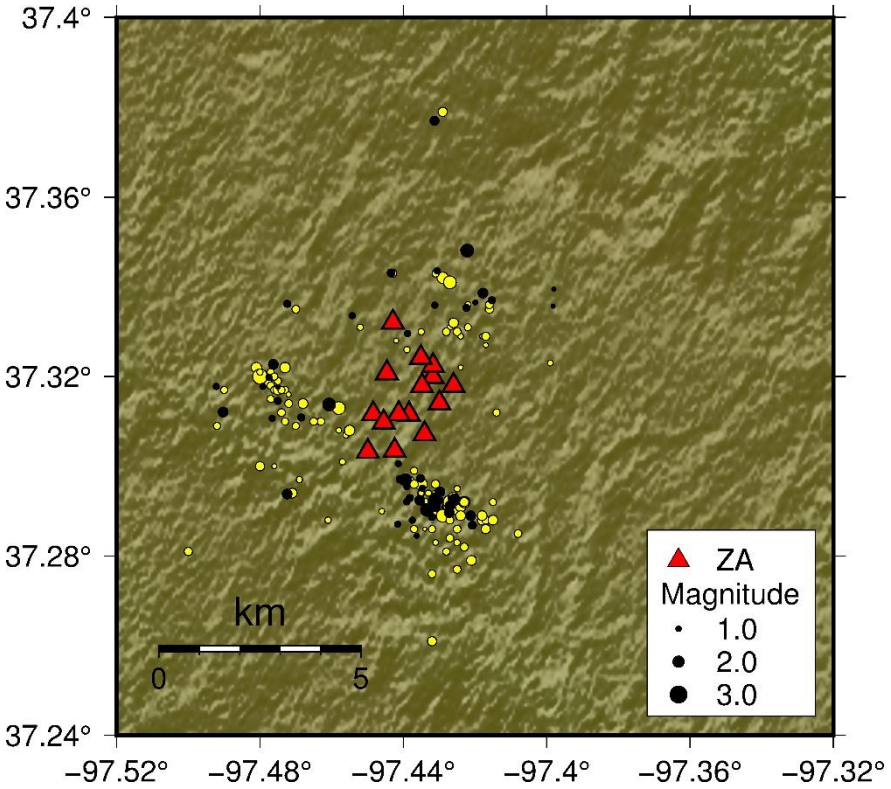

Figure 4.17. Map of the WOF displaying the original catalogue (yellow) and the events that yielded at least one result in the SWS analysis (black). The ZA network is also shown.

In the current thesis, the automatic CA – EV method was employed to analyze waveforms from the Wellington, Kansas CO_2 Sequestration Monitoring (ZA) network (Lynn

Watney, 2014), installed and operated by the Kansas Geological Survey and KU. The network is comprised by twelve broadband stations configured at a sampling rate of 200 sps. The earthquake catalogue (provided by KU) contains 110 events that occurred during 2016 (Figure 4.17). Event magnitudes (M_w) vary between 0.6 and 2.2, while depths are in the range 2.3 – 29.6 km. Finally, the minimum station – event distance is 0.16 km.

4.3.2. Results

For the combination of the CA and EV methods, the initial catalogue was comprised of 392 S – wave arrivals, for the 110 events, out of which 181 met the maximum i_0 criterion. Total processing time was equal to 2 hours and 24 minutes. This catalogue is much more constrained in terms of the angle of incidence, because it was already filtered by the provider.

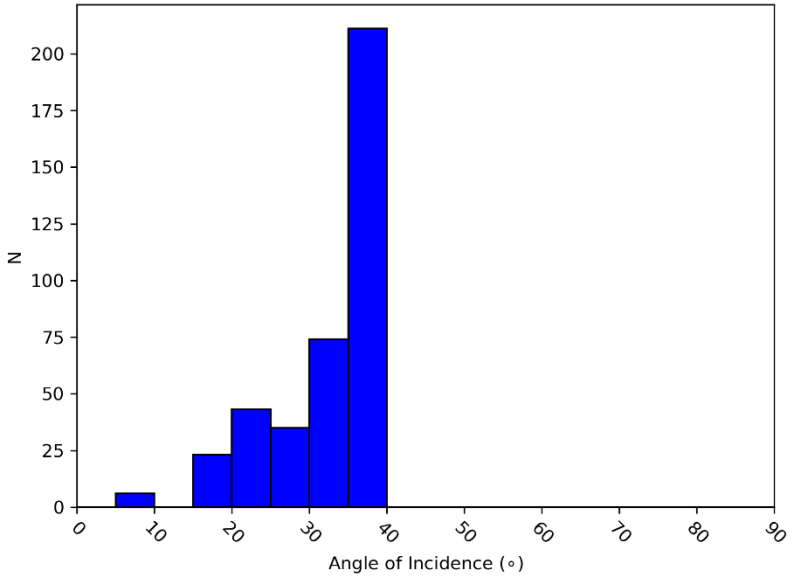

Figure 4.18. Distribution of the angle of incidence in the WOF catalogue. Bin spacing is 5°.

Analysis led to 145 parameter pairs from 49 events. Selection of observations graded between "A" and "C" reduced the number of measurements to 76. Restriction to "A" and "B" grades only (as in the cases of the WGoC and the SVC) would lead to only 36 parameter pairs, rendering the interpretation of the results problematic. In Table 4.4, the summary of the measured parameters is presented for each station. Results are also shown in Figure 4.19.

Table 4.4. Summary of statistics for analyzed SWS parameters in the WOF, with the CA and EV methods. N is the number of observations, $\widetilde{\boldsymbol{\varphi}}$ the median value of the S_{fast} polarization direction, $\boldsymbol{\delta\varphi}$ the standard error of mean of φ, $\overline{\boldsymbol{t_d}}$ the mean time – delay and $\boldsymbol{\delta t_d}$ the standard error of mean of t_d.

Station	Elevation (m)	N	$\widetilde{\boldsymbol{\varphi}}(\boldsymbol{N°E})$	$\boldsymbol{\delta\varphi(N°E)}$	$\overline{\boldsymbol{t_d}}(\boldsymbol{ms})$	$\boldsymbol{\delta t_d(ms)}$
WK01	387	9	173	7	64	13
WK03	384	9	150	3	24	4
WK05	381	8	80	22	79	26
WK06	385	9	17	6	52	10
WK07	385	15	172	6	38	13
WK11	384	14	144	12	39	7

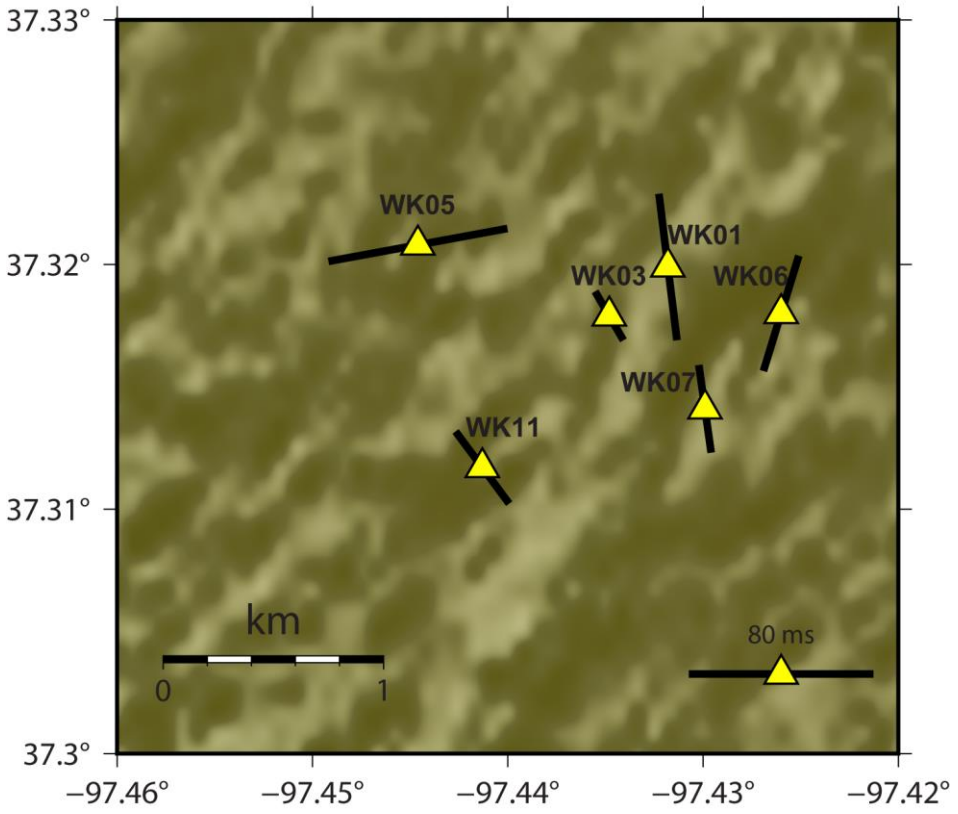

Figure 4.19. Map of the WOF showcasing the median polarization direction for each station. The bar is proportional to the mean time – delay.

Most of the stations seem to present a general NW – SE median polarization direction (Figure 4.21), with the exception of WK05, where $\widetilde{\boldsymbol{\varphi}}$ is oriented WNW – ESE (there seem to be two clear directions, one NE – SW and one NW – SE). It is noted that WK11 presents large scatter of results in virtually all directions. WK03 is the best constrained station. Concerning the time – delays, two groups can be distinguished. Stations WK03, WK07 and WK11 present values equal or lower than 40 ms, while t_d in WK01, WK05, WK06 and WK12 has a bottom limit of 52 ms. There is no correlation between the φ and t_d groups.

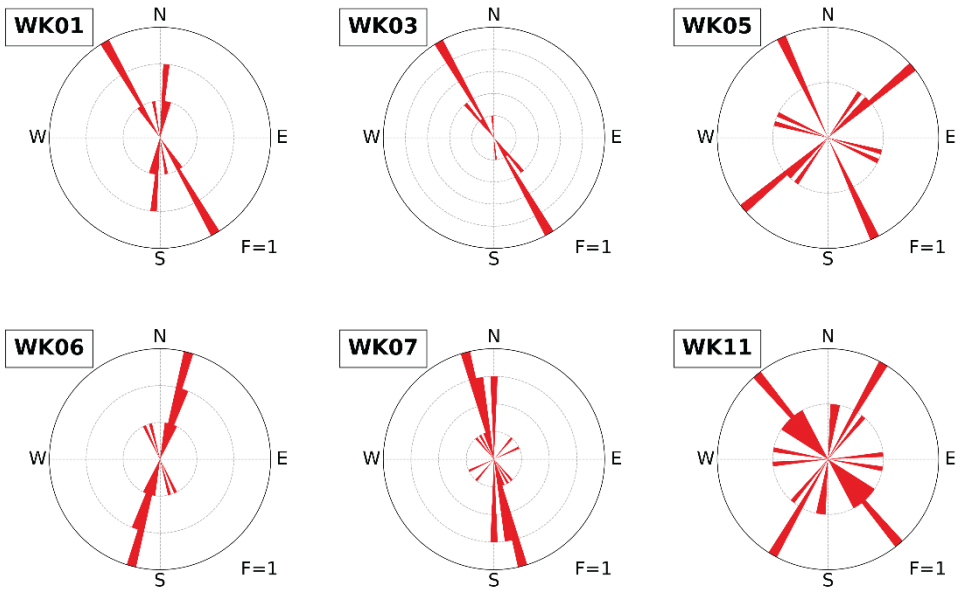

Figure 4.21. Rose diagrams for analyzed SWS parameters in the WOF, with the CA and EV methods. F is the number of observations per grid line.

Considering the proximity of the stations, the individual measurements can be exhibited as a sum. The polarization directions (Figure 4.20) showcase a predominant NW – SE direction. This is in agreement with the results of Nolte et al. (2017) for 2016. The median φ is 168 N°E with a standard error of mean of 4 N°E. Directions are consistent and resemble measurements from a single station. This orientation is in agreement with the observations of flip by Nolte et al. (2017). The histogram of time – delays (Figure 4.20) exhibits a concentration of observations in the 20 – 50 ms range. The mean value of t_d is 45 ± 5 ms.

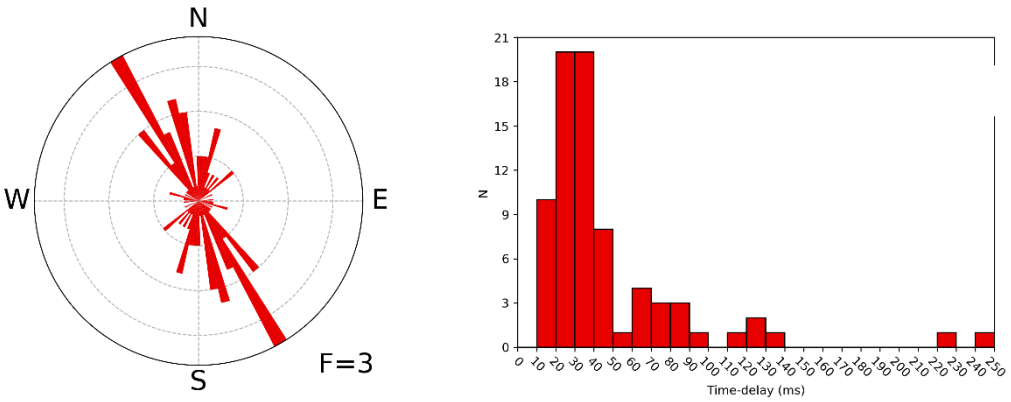

Figure 4.20. Summary plots for analyzed SWS parameters in the WOF, with the CA and EV methods. The rose diagram (left) and the time -delay histogram (right) are presented.

5. Discussion – Conclusions

5.1. Common issues of automatic methods

There are several issues that pose a significant challenge in employing a fully automated solution to study SWS in the upper crust. First and foremost, the selection criteria used are generally very strict. To eliminate the human factor as much as possible, a quantifiable way of evaluating the quality of the waveform must be used. The shear – wave window is a good first – order approach in eliminating secondary phases from the analysis windows. Nevertheless, the ~35° limit obtained from Eq. 2.3 renders a great portion of available data ineligible, even in very active areas, such as the WGoC. In the current study, the 57.7% of available event – station pairs in the area was rejected due to this criterion (see Chapter 4.1.2). Considering the high rejection rate in the next stages of processing for both manual and automatic methods, the shear – wave window can have a devastating effect on a SWS study. Kaviris et al. (2015, 2017, 2018a, 2018d) increased the number of available pairs by extending the window to 45°. To ensure that no secondary phases were erroneously identified as S – arrivals, any recordings presented a higher amplitude in the (initially) determined S – waves in the vertical component, compared to the horizontal ones, were rejected (Kaviris, 2003). A relatively simple clause could be added to the Pytheas software to enforce the above criterion, where windows with higher vertical energy would be rejected. However, identifying converted and scattered phases in local recordings can be challenging, even manually. As the hypocentral distance increases, the classification becomes easier due to phases being separated better (Stephens et al., 1990). Moreover, calculating the angle of incidence of a ray is by itself problematic. No matter the method, there will be uncertainties involved. A good approximation was implemented by using the TauP software and a reliable local 1 – D velocity model. Ray tracing in anisotropic media (Lai et al., 2009; Sadri and Riahi, 2010; Wang, 2014) could facilitate the accurate determination of the angle of incidence.

A second significant problem is the noise content in the waveform and the attempts to remove it. Due to the proximity to the source, as in the case of secondary phases, direct S – arrivals are usually included in the coda of P – waves. This renders the accurate identification of the shear – waves's arrival impossible in some occasions. To reduce the effect of random noise and increase the SNR, a bandpass filter is commonly used. The determination of the proper filter band is crucial to SWS studies. The frequency of the shear – wave is dependent of the source and the distance. Tectonic events featured a lower frequency content compared to volcanotectonic earthquakes or induced seismicity. In addition, events recorded close to the source present higher frequency. To solve this problem, Savage et al. (2010) apply various pre – determined filters to the waveform and select the one that maximizes the SNR as the optimal. However, this process can erroneously choose a filter range that simplifies the pulse and leads to cycle skipping, artificially increasing the time – delay. In the Pytheas software, a different approach has been taken. If the analyst wishes to, they can use the automatic filter determination, where the range is defined as the ±50% of the dominant frequency in the horizontal components. To avoid contamination of long period noise, a minimum cutoff frequency of 1 Hz is also used, meaning that the dominant frequency cannot be less than 1 Hz.

Further research must be conducted on the matter, to develop a more robust method of automatic filtering.

Cycle skipping (whether caused by the filtering or not) is a common issue in SWS studies. Band – limited data and weak signals are the usual causes. In the grid search methods, multiple local minima (for EV) or maxima (for RC) are the indicators of cycle skipping. In Figure 5.1, an example of this issue in SNT1 station of the SVC dataset is shown. Even though this is not included in the final results (SNT1 featured less than ten initial observations), its grade was "D". The rejection of measurements graded "C" and "D" would have excluded it, but for different reasons (high errors). Cycle skipping needs to be manually checked to avoid contamination of the final dataset and artificially high t_d values.

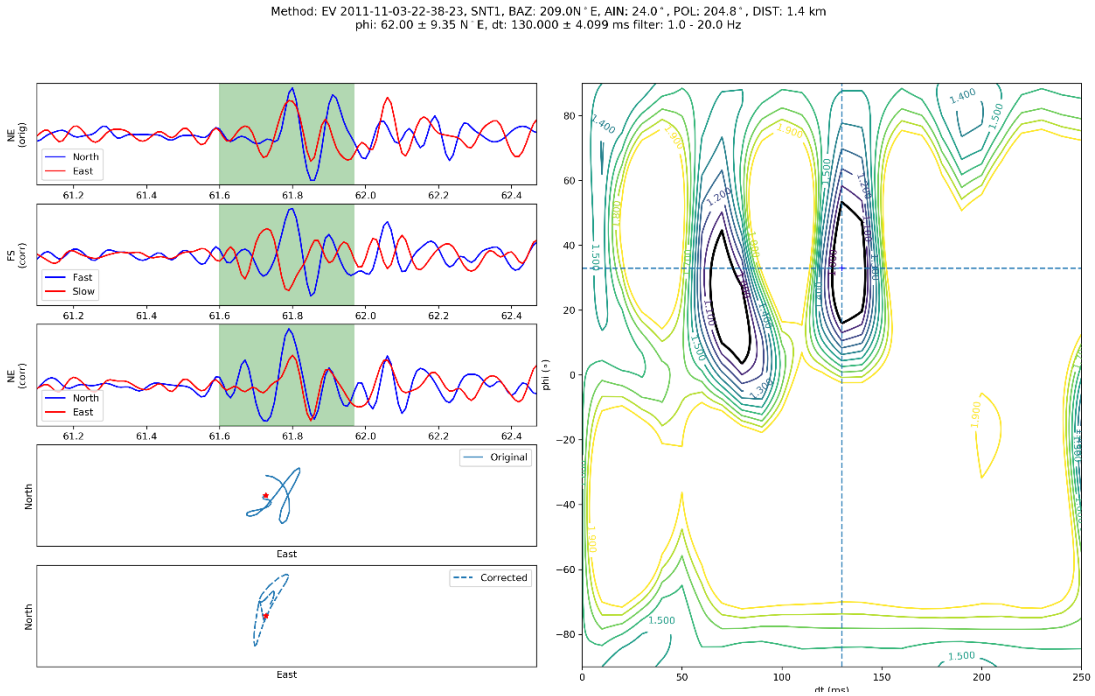

Figure 5.1. Example of cycle skipping from the SVC dataset. The phenomenon is identified by the multiple 95% confidence level contours in the minimum eigenvalue plot. The erroneous matching of the split shear – waves can be seen in the FS rotated system. The S_{slow} has been overly shifted backwards in time. The time – delay from the left 95% confidence region contour (~70 ms) would probably lead to fit data.

All of the SWS methods implemented in Pytheas assume that the split shear – waves are orthogonally polarized. The waveforms are rotated to the FS axial system to apply the time – delay which, for the rotation, is an orthogonal one. There are two types of non – orthogonality: inherent, caused by a strongly anisotropic complex propagation medium (e.g. tilted anisotropy symmetry axis), and apparent, a result of the ray arriving at the station with a non – normal angle of incidence (Li et al., 1998). According to Crampin and Gao (2006), orthogonality is only true for waves that travel with phase velocity. In the current study, an interesting observation arose during data analysis. In the WGoC, stations that provided the most constrained results (for the CA and EV methods) of φ are located in boreholes. Moreover, these measurements do not deviate a lot (Figure 5.2) from results of manual analysis (Kaviris et al., 2018a). This seems to confirm that the interaction of the wave with the free surface can distort the polarization significantly. Since the test for the best fitting φ and t_d is conducted on the corrected waveforms (for either EV or RC), the application of the time – delay on a non - S_{slow} quasi wave will lead to

erroneous parameter estimation. This is a major drawback of grid search methods. Teanby et al. (2004) used data recorded by geophones located in boreholes of 2 km depth. To mitigate the effect of non – orthogonality, a suitable coordinate transformation can be applied to the original data (Lei, 2005).

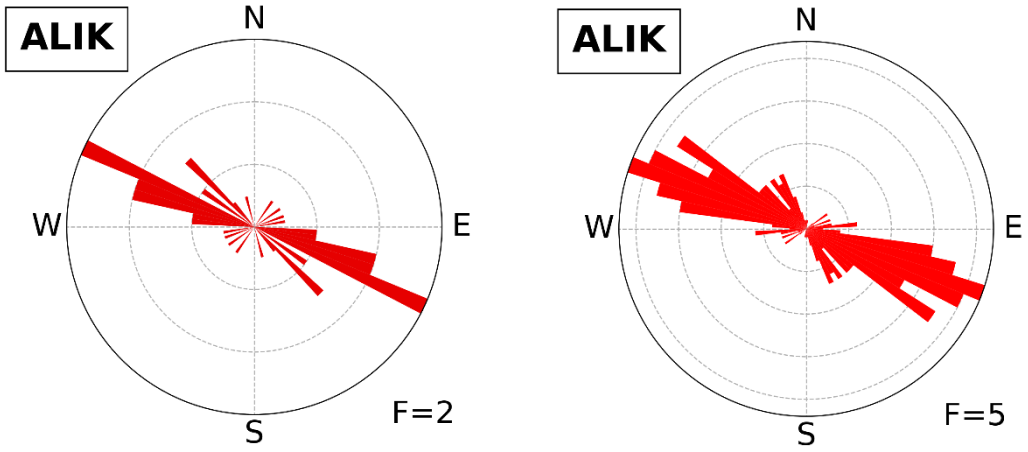

Figure 5.2. Comparison of rose diagrams for the ALIK station (located in a borehole), with measurements obtained from events during 2013 and 2014. The automatic analysis of the current study (left) yielded similar results to the manual analysis (right) of Kaviris et al. (2018a).

5.2. Comparison of automatic and manual techniques

To investigate the reliability of automatic measurements, they were compared to ones obtained from the manual analysis of Kaviris et al.(2017) for the WGoC (during the first semester of 2013) and Kaviris et al.(2015) for the SVC. Only graded "A" and "B" observations (in both manual and automatic catalogues) were used in the comparison. The correlation coefficient (R) for φ values was obtained from Jammalamadaka and SenGupta (2001) to accommodate wrapping phenomena at the edges of the angle range. For the WGoC dataset (Figure 5.3), the φ values show good correlation between them, with $R = 0.61$. On the other hand, time – delays showcase adequate correlation visually, but the Pearson's correlation coefficient ($R = -0.26$) does not substantiate this observation, statistically.

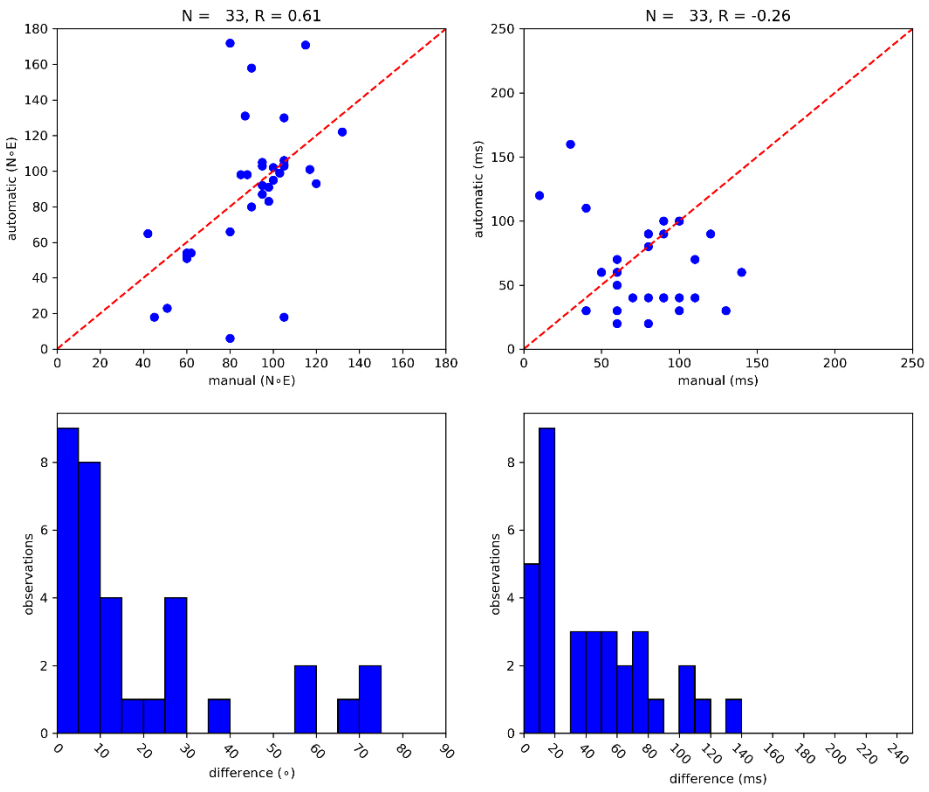

Figure 5.3. Comparison plot for measurements of automatic and manual catalogues in the WGoC. The correlation plot of φ (top left) shows multiple measurements close to the linear model $\boldsymbol{x} = \boldsymbol{y}$ (dashed line). This is also shown in the difference histogram (bottom left). Time – delays (top right) are more scattered and cannot provide a reliable $\boldsymbol{R}$. It is noted that the most populated class in the histogram (bottom right) are the ones that represent a maximum difference of 20 ms (i.e. 2 samples). N is the number of observations and $\boldsymbol{R}$ the correlation coefficient.

The difference between the median S_{fast} polarization directions obtained from the automatic analysis and the average ones from Kaviris et al. (2018d) are not so pronounced (Figure 5.4). The residuals between manual and automatic observations vary from zero in LAKA to 42° in KALE. Three stations (KALE, MG00 and PYRG) are of special interest. Automatic measurements vary greatly (over 35°) in these. KALE presents high scattering. In PYRG, the manually determined mean φ (56°) can be seen as a secondary direction (Figure 4.3).

However, an interesting comparison is present in MG00. For the automatic method, analyzed events were constrained between January and June of 2013, while the Kaviris et al. (2018d) dataset contains station – event pairs for MG00 mainly in 2014. Could this describe a potential change in direction? This hypothesis must be investigated by measuring SWS in the station for 2014 with the automatic method. If the results are reliable and the new direction is similar to the manual one, then polarization shifted. Regardless, the direction obtained in the current study agrees with nearby MALA and SERG, forming a larger region of NE – SW dominant S_{fast} axis strike. The median value of φ in EFP showcases a similar direction, but the station is marred by very high scatter in measurements and the results is not reliable. Differences in mean time – delays are not discussed, due to their volatile characteristics. A comparison of this parameter on a sample – to – sample basis is more valid (as seen above).

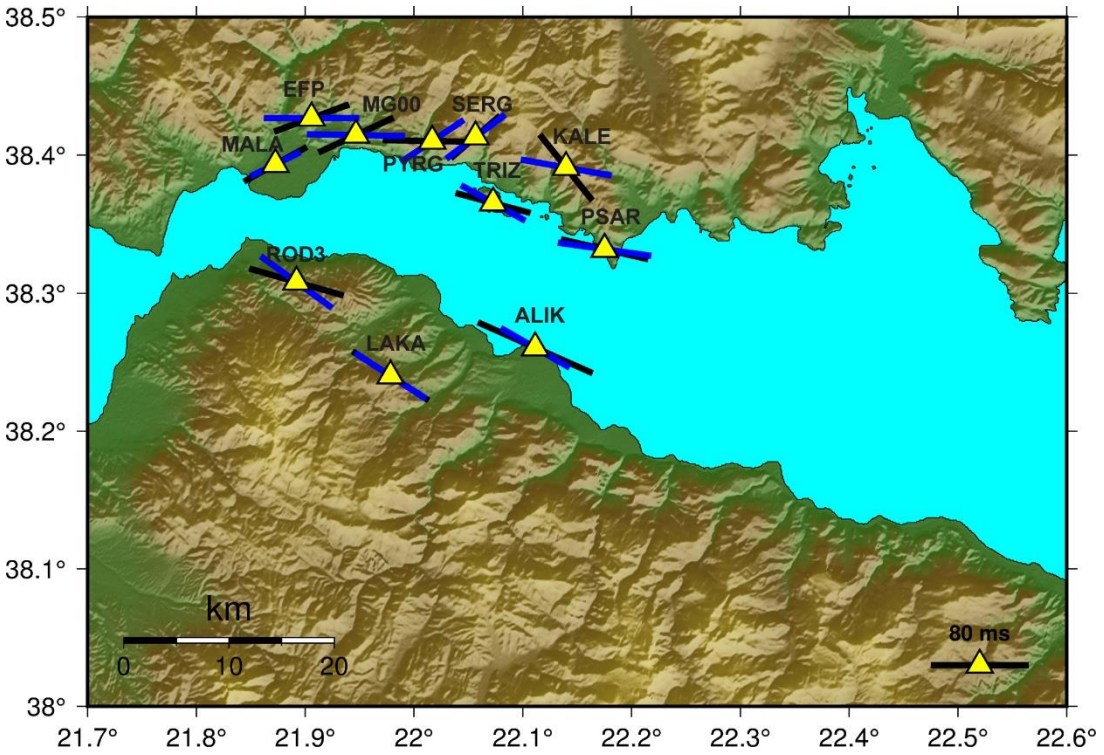

Figure 5.4. Map of shear – wave splitting results in the WGoC from the automatic method (black bars). The results of Kaviris et al. (2018b) are overlaid (blue bars).

The opposite behavior Is observed in the SVC data (Figure 5.5), which are compared to the manual ones from Kaviris et al. (2015). t_d values have $R = 0.71$, while polarization directions do not seem to be statistically correlated ($R = 0.36$). This is also evident from the histograms, where time – delays differences are largely limited to up to 50 ms (5 samples). Nevertheless, by visually inspecting the diagrams it is evident that there is a significant number of automatic measurements that match the manual ones. Deviations and outliers are attributed to badly graded measurements in the CA catalogue and human bias in manual analysis.

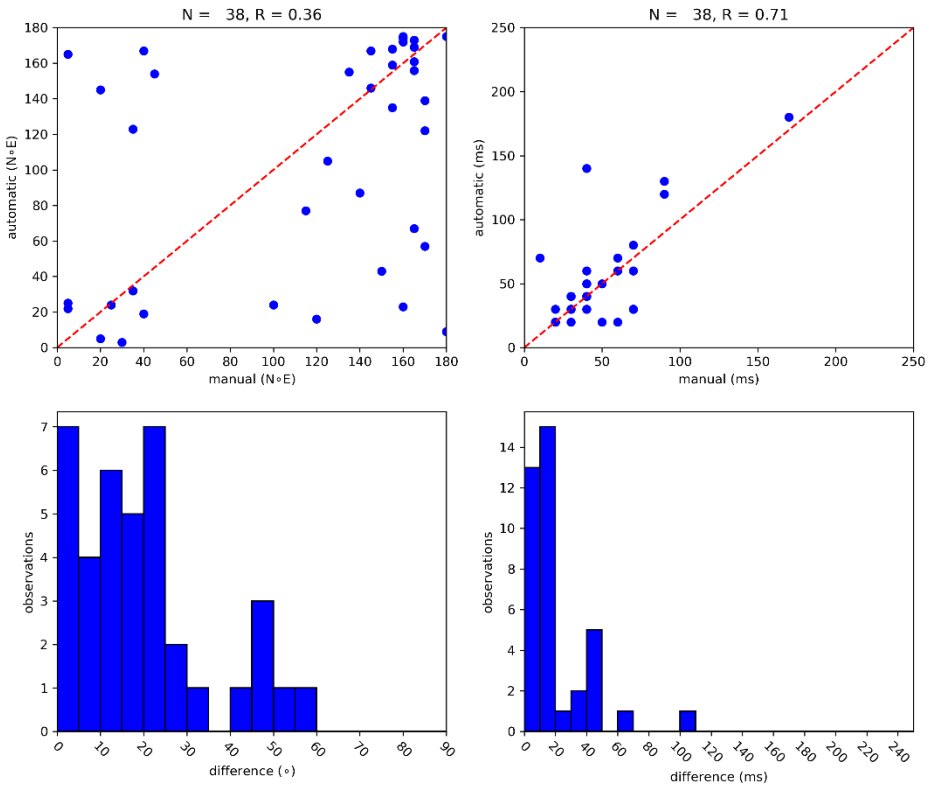

Figure 5.5. Comparison plot for measurements of automatic and manual catalogues in the WGoC. The correlation plot of φ (top left) shows multiple measurements close to the linear model $\boldsymbol{x} = \boldsymbol{y}$ (dashed line). This is also shown in the differences histogram (bottom left). Time – delays (top right) are approximating the line and are characterized by a high $\boldsymbol{R}$. It is noted that the most populated classes in the histogram (bottom right) are the ones that represent a maximum difference of 20 ms (i.e. 2 samples). N is the number of observations and $\boldsymbol{R}$ the correlation coefficient.

A comparison between the average parameters from the manual and automatic processes yields optimistic results. THT1, the station with most observations in either study, has virtually the same values. The average φ differs only by 1° and even the variation in t_d is rather small (7 ms). On the contrary, SNT3 showcases a 23° discrepancy in S_{fast} directions and 26 ms in mean t_d. SANT yielded only one result in the final automatic dataset and, as such, no conclusion can be drawn. At this point, it is important to mention another major difference between the manual and the automatic results for the SVC. The study of Kaviris et al. (2015) includes observations in eight stations, for a total of 340 parameter pairs. Stations CMBO and SAP3 are located further away from the seismicity cluster, leading to angles of incidence outside

of the shear – wave window defined in the current study. The manual analysis was conducted for maximum angle of incidence equal to 45°, instead of 35°, thus excluding these two stations. The same is valid for SANT, but one observation (corresponding to an event that occurred really close to the station, about 1.5 km to the west) managed to find its way into the final dataset. However, the absence of results for stations SAP2, SNT1 and THT2 is troubling. Further investigation led to the discovery of results in all three in the initial catalogue. However, SAP2 and SNT1 were rejected due to the minimum number of ten results criterion. The former yielded only 5 observations (instead of the 89 from the manual analysis) and the latter 7 (instead of 8). Considering that SNT3 (who is located next to SAP2) featured an initial number of 28 observations, processing in this station needs to be investigated further. The initial 12 results in THT2 were rejected due to grading ("C" and below).

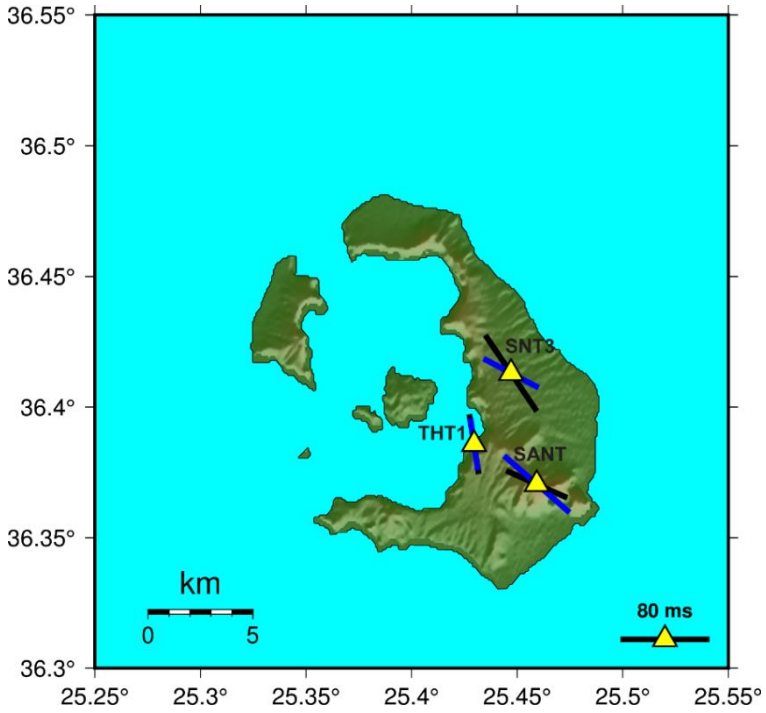

Figure 5.6. Map of shear – wave splitting results in the SVC from the automatic method (black bars). The results of Kaviris et al. (2015) are overlaid (blue bars).

Finally, a potential difference in the number of results may stem from the event location process. In the current study, hypocenters were acquired from Hypoinverse. In both Kaviris et al. (2018b) and Kaviris et al. (2015) a relocation was conducted through a double – difference algorithm (Waldhauser, 2001) and determined more accurate results. This process has surely led to discrepancies between the event databases.

In conclusion, the comparison between manually and automatically acquired results provided valuable insight on the applicability of the CA and EV methods. Concerning the polarization direction of the S_{fast}, the automatic method seems satisfyingly reliable. With the exception of some outliers (e.g. stations PYRG and KALE) observations between methods are similar. The same could be stated for the time – delays, although the comparison yields slightly worse results. Further work in the quality assessment and grading of automatic measurements could improve the reliability of the method. Nevertheless, human error and bias can considered as an important factor in discrepancies between the two techniques.

5.3. Assessment of the Grading Algorithm

Grading could well be considered the second most important part of an automatic processing scheme, after the analysis itself. A robust grading algorithm assigns weights to the measurements and facilitates the automatic sorting of observations in categories. Quantifying the quality criteria is a difficult task. To do so, bounds for specific parameters must be set. However, there is the issue of how these bounds are selected. The expertise of the analyst has an important role in this, as these must be manually defined, usually from experience. As described in further detail in Chapter 3.3.1, Pytheas incorporates the comparison of three parameters to automatically evaluate the quality of the measurement: (i) the error of the φ, (ii) the error of the t_d and (iii) the SNR. The limits were selected (Chapter 3.4) after inspecting a sample number of measurements.

An example of the grading for the SVC dataset is provided below (Figure 5.7). Grading determined 6 null measurements, out of the 145 in the input. The majority of the rest (82), though, was sorted in the "A" and "B" categories. The "E" grade (i.e. at least one of the three grading criteria was not fulfilled) was assigned to 37 values, where the low SNR was the major factor of rejection.

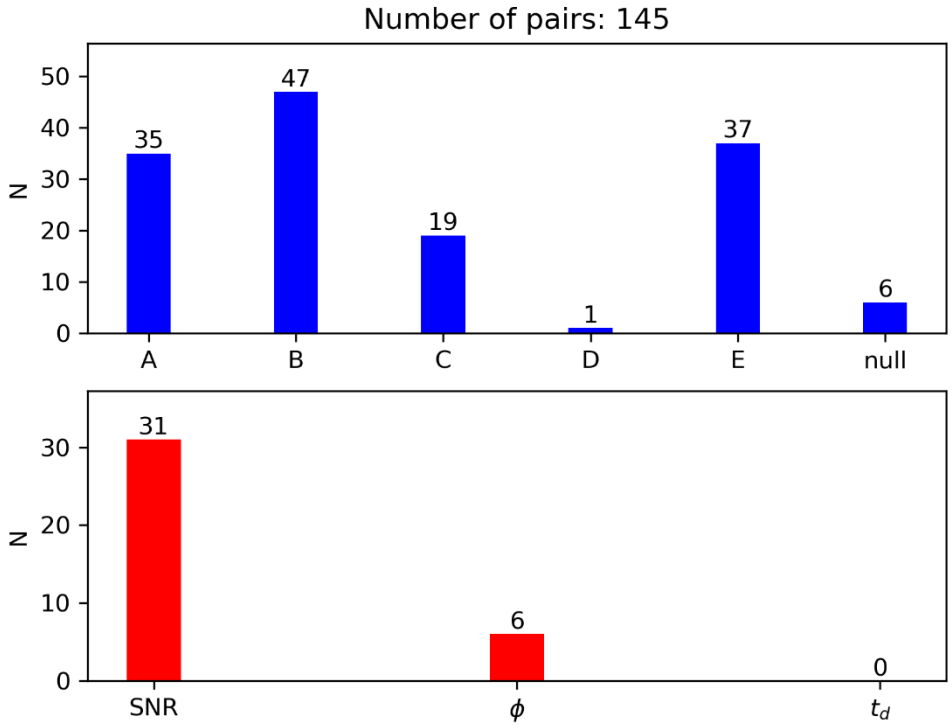

Figure 5.7. Distribution of grades (top) and rejection causes (bottom) for the grading procedure. N is the number of observations (also located on top of each bar).

To assess the performance of the algorithm, the variation of the standard deviation of φ (σ_φ) and t_d (σ_{t_d}) in all three SWS catalogues (WGoC, SVC and WOF) in relation to the minimum accepted grade is investigated. First, it has to be considered that, by applying stricter criteria (i.e. the grade going from "E" to "A") the population of the samples decreases, affecting the standard deviation. Of the 22 stations presented in the current study, only two included over 10 observations in the "A" class. For those two (AIOA and THT1), deviation decreased consistently with the increase in grading. σ_φ decreased by 35.8% and σ_{t_d} by 62.4%, from "E" to "A", in AIOA station. A similar behavior was observed in THT1, where σ_φ saw a 22.2% reduction and σ_{t_d} decreased by 37.8%. The number of observations declined by 60.5% and 67.8%, respectively. For the rest of the stations, the results are summarized in Table 5.1.

Table 5.1. Perturbations of number of observations ($\boldsymbol{dN}$), standard deviation of φ ($\boldsymbol{d\sigma_\varphi}$) and t_d ($\mathbf{d}\boldsymbol{\sigma_{t_d}}$) between grade classes "E" and "B". For the stations of the WOF (in italics), classes "E" and "C" were used. "nan" values refer to the inability of calculation the standard deviation due to limitations in the number of observations.

Station	$\boldsymbol{dN}$ (%)	$\boldsymbol{d\sigma_\varphi}$ (%)	$\mathbf{d}\boldsymbol{\sigma_{t_d}}$ (%)
AIOA	-23.3	-14.1	-44.6
ALIK	-72.6	-26.5	-16.4
EFP	-54.2	-18.8	-16.0
KALE	-64.0	16.2	-26.3
LAKA	-75.8	-22.0	-27.8
MALA	-68.2	-19.1	-24.1
MG00	-38.5	-10.3	-27.7
PANR	-64.7	-45.0	4.8
PSAR	-63.5	-37.7	-22.2
PYRG	-62.5	-7.8	-6.0
ROD3	-66.7	-56.6	-8.3
SANT	-93.3	nan	nan
SERG	-54.9	-28.2	-28.0
SNT3	-46.4	-12.6	-2.1
THT1	-30.0	-7.9	-15.5
TRIZ	-71.1	-10.3	-20.2
WK01	-35.7	13.2	-6.5
WK03	-50.0	-44.0	-87.8
WK05	-27.3	39.0	-36.5
WK06	-30.8	-36.3	-4.2
WK07	-50.0	-21.5	-84.9
WK11	-56.2	8.3	-63.5

It is evident that the grading algorithm successfully removed the bulk of spurious measurements from the dataset. There are some exceptions concerning φ, in which case the standard deviation increased (KALE, WK01, WK05 and WK11). The large rejection rate of measurements is not uncommon in SWS studies, automatic or not (Teanby et al., 2004; Savage et al., 2010; Al-Harrasi et al., 2011; Kaviris et al., 2017, 2018a, 2018b). In the current study, the average rejection rate was 54.5%. The bounding values used in the grading should be examined and potential revisions could result from more experience in using the algorithm.

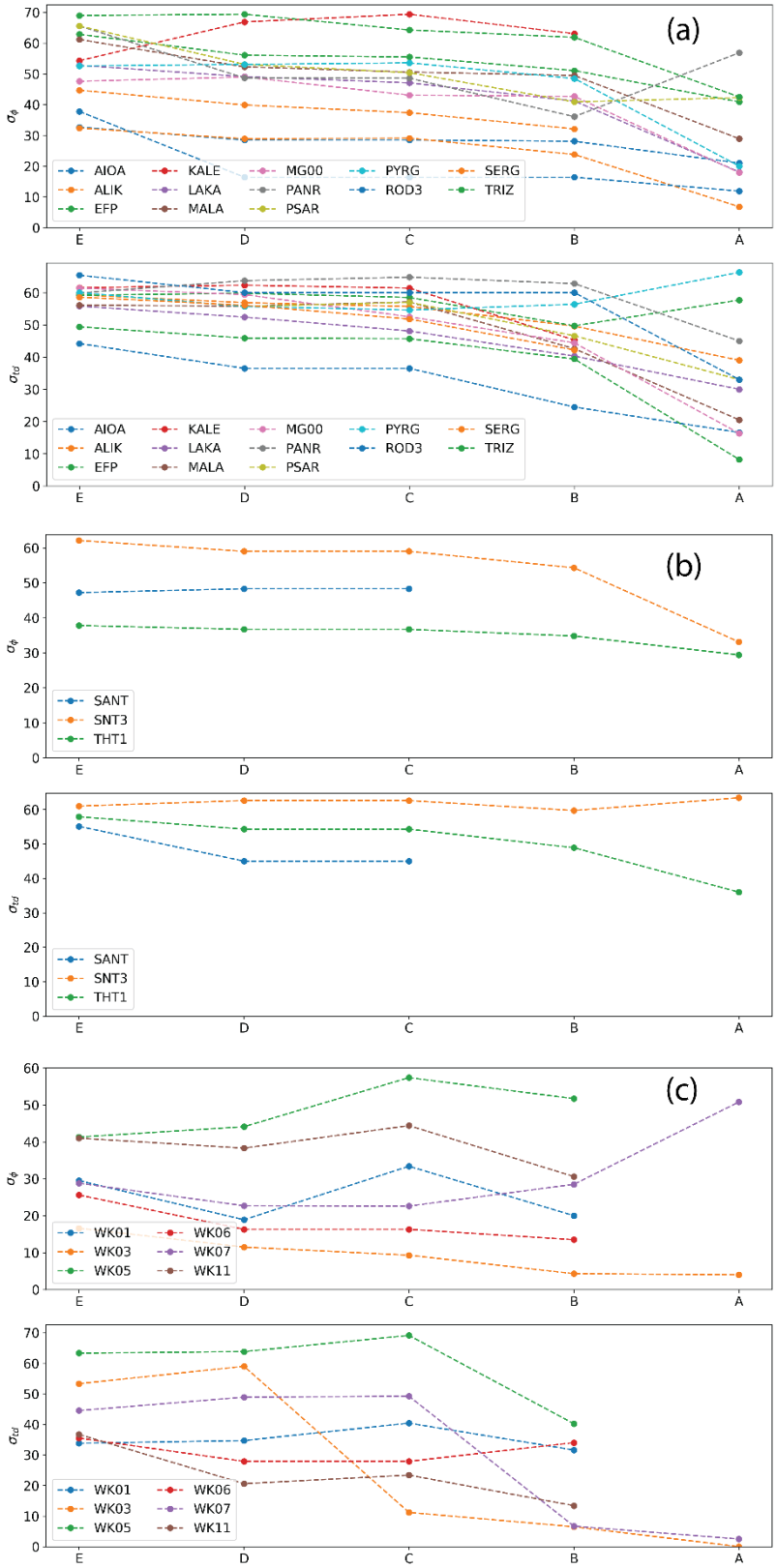

Figure 5.8. Variations of standard deviation for φ (top plots) and t_d (bottom) for the WGoC (a), SVC (b) and WOF(c), according to the minimum accepted grade. Classes without a data point did not have any available results. There is a consistent reduction observed, interspersed by spikes of increase related to the decrease in the number of observations. This behavior is more pronounced in stations with a small amount of initial measurements (i.e. the ones in the WOF) and in the "A" class (where the vast majority of parameters is usually rejected).

5.4. On the Pytheas software

The main advantage of Pytheas is the seamless integration of the most popular SWS techniques under one GUI. This enables the user to select the analysis scheme *ad – hoc* and easily compare between methods. Furthermore, results from (semi- or not) automatic methods can be effortlessly verified in a couple of seconds. For example, if a user wants to apply the CA and EV methods on a single station – event pair, they have to follow these steps: (1) pick the S_{fast} – arrival (if not already provided by the catalogue), (2) select the CA – EV method from the splitting options (or use the Ctrl+T hotkey) and (3), if they wish, rotate waveforms to the corrected NE system (Ctrl+3 hotkey) to verify the result by inspecting the traces and both particle motion diagrams (polarigram and hodogram). This way, the user can quickly evaluate the measurements of both φ and t_d. The starting screen of each station – event pair displays the polarigram. Upon picking the S_{fast} – arrival, the user instantly has determined φ manually. After steps (1) and (2), the linearity of the polarigram vectors at the (now) S – arrival is an indication on the quality of the t_d. On a modern computer, the CA – EV processing time for one example is about a minute (dependent of the number of windows). This means that, including manual verification, the analysis shouldn't take over two minutes, even for an inexperienced user. This is a major advantage that will facilitate increasing the robustness of results, as well as render the analysis more effective. If, instead of a pair – to – pair basis, the user wishes to utilize the catalogue – wide automatic processing, opening (afterwards) the event – station pair will load the results of the processing. Hence, performing a manual quality control check on the fully automatic scheme is also time – effective.

The current author has participated in the manual analysis of three SWS datasets (Kaviris et al., 2017, 2018a, 2018b). With the original (unpublished) POLAR software used (written in Fortran and Shell scripts), each event – station pair would take about 8 minutes to process (even after years of experience). With the use of Pytheas, the manual analysis time was reduced to ~3 minutes, a 62.5% decrease. In addition, this is currently the only freely available software that incorporates the PM method.

The picking of S – arrivals from routine analysis is not always suitable for a SWS study. This is critical for the application of the automatic methods. It is partially mitigated by the configuration of windowing in CA, by selecting a minimum window that allows for an acceptable residual. Nevertheless, Pytheas can be used to pick the S_{fast}. The shear – wave window limit can be set globally. As a result, selecting to go to the next station or event (Ctrl+W and Shift+N, respectively) will choose the next pair that falls within it. Picking is, thus, rendered rapid, with multiple arrivals be able to be determined in swift succession. Moreover, to facilitate this, the user can choose to take into account the result of the automatic AR – AIC picker and the theoretical arrival time calculated by TauP. Concerning the arrivals, the optional use of TauP can generate more accurate angles of incidence than the ones in the catalogue, provided a suitable velocity model is given. This is crucial for SWS analysis, as i_0 is used in both selecting rays and rotating waveforms to the LQT system.

The requirements to use the software are standardized. The dependencies are comprised by popular Python packages, which can be easily installed through a package manager (e.g. PyPi or Anaconda). The configuration of the application's components can be

achieved through modifying easily readable ASCII files. The input catalogue format is QuakeML. This permits the standardization of event information and does not require the user to develop additional tools to convert their catalogue to a niche format. Admittedly, it is not always an easy task to produce a QuakeML file (at least for users that are not familiar with programming). However, its increasing use in the seismological community should eradicate this issue in time. Online catalogue services of the FDSN already offer the choice of downloading event data in a QuakeML format (including arrivals and picks). The format of the waveforms is also very accessible. Any trace file compatible with Obspy can be used (SAC, MSEED, etc.). There are a few details that need to be given extra care. First, due to design choice, Pytheas does not merge data. This is to avoid erroneous merges or conceal underlying issues with the waveforms. Thus, the provided waveforms must not be segmented. The trace files do not have to meet any specific naming convention, as long as their name includes the station code and they are contained in a folder named after the event code. This event code will be used to match event data to waveforms.

In conclusion, Pytheas is focused on the user accessibility aspect, without sacrificing anything in terms of processing robustness and effectiveness. There are several areas of improvement, for subsequent releases of the software. The implementation of a more robust auto – picker could aid in requiring only the event and waveform data for the analysis, without having any arrivals information and not necessitating the manual picking by the user. Furthermore, a more reliable method of automatically determining the optimal filter frequencies needs to be implemented. While the broad 1 – 20 Hz filter used in the current thesis yielded satisfactory results, there are situations where a stricter band could have removed more of the noise content without distorting the shear – wave signal. Grading needs to incorporate more metrics, such as waveform similarity of the corrected traces and an estimate of the energy content in the corrected transverse component (for EV). Moreover, there is the question of determining the limits of parameters for grading. The empirically obtained values need to be further refinement to permit the inclusion of more quality observations in the final dataset. Another section where Pytheas could improve on is results visualization and processing. The statistics and related figures (rose diagrams, maps, histograms) can be generated automatically at the end of the analysis. This will be a significant QoL feature, as it will not require the user to develop their own tools for the above operations. Similarly, functionality concerning temporal variations of time -delays and polarization directions, such as moving averages (Crampin et al., 1999), normalized t_d error estimation (Del Pezzo et al., 2004) and sorting of observations in bands according to the incident ray and the microcrack orientation (Gao and Crampin, 2006; Crampin, 2011; Crampin et al., 2013). The above is used in monitoring stress variations and can lead to a possible "stress – forecast" of an earthquake or volcanic event (Gao and Crampin, 2004; Bianco et al., 2006; Crampin and Gao, 2012; Crampin et al., 2015). Finally, the Obspy package offers a multitude of routines to access online repositories of seismological information. Incorporating these to Pytheas will permit the user to obtain both event and waveform data easily, in available areas. While this can already be done quite easily. However, Pytheas would offer a more SWS – centric approach, for example, with data selection based on a given shear – wave window.

5.5. Applicability of the Shear – Wave Splitting Tomography

The application of the SWST method led to valuable insight concerning its application eligibility. The fundamental issue with the technique is one that prevalent in SWS studies. As with every tomographic inversion, SWST's reliability is heavily dependent on the ray coverage. However, the high rejection rate of observations severely limits the available wave paths. This issue is more conspicuous in automatic methods. In the current thesis, results from CA and EV could not be used due to their limited number. They could not yield any reliable results, by consistently failing the checkerboard tests. Employing manual measurements by previous studies managed to produce areas of resolution. However, in the SVC the resolution contrast with the WGoC (where the number of observations was almost ten times higher) is striking. The utilization of wider grid cells (in respect to each region's geographical extent) led to lower resolution and worse defined checkerboards.

Another issue is related with the concept of time – delay. The assumption that it is accumulated along the ray path produces artificially high anisotropy at station locations. Moreover, the eligibility of using time – delays over long periods needs to be questioned. The t_d is a highly volatile quantity with daily changes, as a response to the medium's stress state. It is only natural that incorporating measurements over years will not yield a comprehensive image of anisotropy in the crust. SWST should be conducted in time – intervals, which has the disadvantage of greatly reducing the number of measurements per inversion. To determine these periods, the occurrence of significant tectonic or volcanic events can be used to outline expected stress variations. In areas where induced seismicity is dominant, these time – lapses can be determined more effectively, by changes in the earthquake cause (e.g. alterations in oil production or specific water injections). An additional problem with time – delays is their scatter. As detailed in Chapter 1.2, t_d scattering is a common observation in SWS studies. The model will have to fit, for similar ray paths, significantly different measurements. This will result in a dubious distribution of anisotropy in the medium.

To successfully interpret a tomography obtained through SWS, further information is required. As explained in the above paragraph, for areas such as the ones examined in the current thesis, anisotropy is directly affected by stress. Thus, other methods need to be used in tandem with SWST, to prove its reliability. For example, pressure measurements or inversion of focal – mechanisms (Kassaras and Kapetanidis, 2018) could be used to directly be compared with results of the tomography. In the SVC, there are two structures observed in the resolution area, striking approximately N – S, between Santorini and Nea Kameni (Figure 4.16). These are sub – parallel to low S- velocity shallow anomalies obtained from local tomography (Papadimitriou et al., 2015) and attributed to activation of the magmatic chamber. For the WGoC, there are alternating moderate and low anisotropy regions (Figure 4.8), which do not form a pattern. Even though the checkerboard tests define a well – resolved area, the variability of time – delays due to the long analysis period (two years) has probably distorted any reliable results. From the above, it is clear the method of SWST requires further investigation before it can be considered reliable. The main issue, however, is the data availability. Given enough measurements, a 2D inversion (in determined time intervals) could produce critical information on the lateral variations of stress in an anisotropic medium.

References

Akazawa, T., 2004. A Technique for Automatic Detection of Onset Time of P- and S-Phases in Strong Motion Records. 13th World Conf. Earthq. Eng. 786.

Al-Harrasi, O.H., Al-Anboori, A., Wüstefeld, A., Kendall, J.-M., 2010. Seismic anisotropy in a hydrocarbon field estimated from microseismic data. Geophys. Prospect. 59, 227–243. doi: 10.1111/j.1365-2478.2010.00915.x

Al-Harrasi, O.H., Kendall, J.M., Chapman, M., 2011. Fracture characterization using frequency-dependent shear wave anisotropy analysis of microseismic data. Geophys. J. Int. 185, 1059–1070. doi: 10.1111/j.1365-246X.2011.04997.x

Ambraseys, N.N., Jackson, J.A., 1990. Seismicity and associated strain of central Greece between 1890 and 1988. Geophys. J. Int. 101, 663–708.

Armijo, R., Meyer, B., King, G., Rigo, A., Papanastassiou, D., 1996. Quaternary evolution of the Corinth Rift and its implications for the Late Cenozoic evolution of the Aegean. Geophys. J. Int. 11–53.

Aster, R.C., Shearer, P.M., Berger, J., 1990. Quantitative measurements of shear wave polarizations at the Anza Seismic Network, southern California: Implications for shear wave splitting and earthquake prediction. J. Geophys. Res. 95, 12449. doi: 10.1029/JB095iB08p12449

Avallone, A., Briole, P., Agatza-Balodimou, A.M., Billiris, H., Charade, O., Mitsakaki, C., Nercessian, A., Papazissi, K., Paradissis, D., Veis, G., 2004. Analysis of eleven years of deformation measured by GPS in the Corinth Rift Laboratory area. Comptes Rendus - Geosci. 336, 301–311. doi: 10.1016/j.crte.2003.12.007

Baccheschi, P., Margheriti, L., Steckler, M.S., 2008. SKS splitting in Southern Italy: Anisotropy variations in a fragmented subduction zone. Tectonophysics 462, 49–67. doi: 10.1016/j.tecto.2007.10.014

Bale, R., Gratacos, B., Mattocks, B., Roche, S., Popvlaskii, K., Xinxiang, L., 2009. Shear wave splitting applications for fracture analysis and improved imaging: Some onshore examples. First Break 27, 73–83.

Bell, R.E., McNeill, L.C., Bull, J.M., Henstock, T.J., 2008. Evolution of the offshore western Gulf of Corinth. Bull. Geol. Soc. Am. 120, 156–178. doi: 10.1130/B26212.1

Berens, P., 2009. CircStat : A *MATLAB* Toolbox for Circular Statistics. J. Stat. Softw. 31. doi: 10.18637/jss.v031.i10

Bernard, P., Briole, P., Meyer, B., Gomez, J., Tiberi, C., Berge, C., Cattin, R., Hatzfeld, D., Lachet, C., Lebrun, B., Deschamps, A., Courboulex, F., Larroque, C., Rigo, A., Massonnet, D., Papadimitriou, P., Kassaras, J., Diagourtas, D., Makropoulos, K., Veis, G., Papazisi, E., Mitsakaki, C., Karakostas, V., Papadimitriou, E., 1997. The Ms=6.2, June 15, 1995 Aigion earthquake (Greece): evidence for low-angle normal faulting in the Corinth rift. J. Seismol. 1, 131–150. doi: 10.1023/A:1009795618839

Bernard, P., Zollo, A., 1989. Inversion of near-source S polarization for parameters of double-couple point sources. Bull. Seismol. Soc. Am. 79, 1779–1809.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., Wassermann, J., 2010. ObsPy: A Python Toolbox for Seismology. Seismol. Res. Lett. 81, 530 LP-533.

Bianco, F., Scarfì, L., Del Pezzo, E., Patanè, D., 2006. Shear wave splitting changes associated with the 2001 volcanic eruption on Mt Etna. Geophys. J. Int. 167, 959–967. doi: 10.1111/j.1365-246X.2006.03152.x

Bianco, F., Zaccarelli, L., 2009. A reappraisal of shear wave splitting parameters from Italian active volcanic areas through a semiautomatic algorithm. J. Seismol. 13, 253–266. doi: 10.1007/s10950-008-9125-z

Bohnhoff, M., Rische, M., Meier, T., Becker, D., Stavrakakis, G., Harjes, H.P., 2006. Microseismic activity in the Hellenic Volcanic Arc, Greece, with emphasis on the seismotectonic setting of the Santorini-Amorgos zone. Tectonophysics 423, 17–33. doi: 10.1016/j.tecto.2006.03.024

Booth, D., Crampin, S., 1985. Shear-wave polarizations on a curved wavefront at an isotropic free surface. Geophys. J. Int. 31–45.

Bouin, M.-P., Téllez, J., Bernard, P., 1996. Seismic anisotropy around the Gulf of Corinth, Greece, deduced from three-component seismograms of local earthquakes and its relationship with crustal strain. J. Geophys. Res. Solid Earth 101, 5797–5811. doi: 10.1029/95JB03464

Bowman, J., Ando, M., 1987. Shear-wave splitting in the upper-mantle wedge above the Tonga subduction zone. Geophys. J. Int. 25–41.

Briole, P., Avallone, A., Billiris, H., Charade, O., 2003. A ten years analysis of deformation in the Corinthian Gulf via GPS and SAR Interferometry. 11th Int. Symp. Deform. Meas.

Brooks, M., Ferentinos, G., 1984. Tectonics and sedimentation in the Gulf of Corinth and the Zakynthos and Kefallinia channels, Western Greece. Tectonophysics 101, 25–54. doi: 10.1016/0040-1951(84)90040-4

Brooks, S.G., Chroston, P.N., Booth, D.C., 1987. Extensive-dilatancy anisotropy (EDA) inferred from observations of crustal shear waves generated by a refraction experiment in northern Scandinavia. Geophys. J. R. Astonomical Soc. 225–232.

Caliński, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Commun. Stat. 3, 1–27. doi: 10.1080/03610927408827101

Chouliaras, G., Kassaras, I., Kapetanidis, V., Petrou, P., Drakatos, G., 2015. Seismotectonic analysis of the 2013 seismic sequence at the western Corinth Rift. J. Geodyn. 90, 42–57. doi: 10.1016/j.jog.2015.07.001

Crampin, S., 2011. Shear-Wave Splitting: New Geophysics and Earthquake Stress-Forecasting, in: Gupta, H.K. (Ed.), Encyclopedia of Solid Earth Geophysics. Springer, Dordrecht. doi: 10.1007/978-90-481-8702-7

Crampin, S., 1984. An introduction to wave propagation in anisotropic media. Geophys. J. Int. 76, 17-28. doi: 10.1111/j.1365-246X.1984.tb05018.x

Crampin, S., 1978. Seismic-wave propagation through a cracked solid: polarization as a possible dilatancy diagnostic. Geophys. J. Int. 53, 467–496. doi: 10.1111/j.1365-246X.1978.tb03754.x

Crampin, S., Evans, R., Atkinson, B.K., 1984. Earthquake prediction: a new physical basis. Geophys. J. R. Astron. Soc. 76, 147–156.

Crampin, S., Gao, Y., 2012. Plate-wide deformation before the Sumatra-Andaman Earthquake. J. Asian Earth Sci. 46, 61–69. doi: 10.1016/j.jseaes.2011.10.015

Crampin, S., Gao, Y., 2006. A review of techniques for measuring shear-wave splitting above small earthquakes. Phys. Earth Planet. Inter. 159, 1–14. doi: 10.1016/j.pepi.2006.06.002

Crampin, S., Gao, Y., Bukits, J., 2015. A review of retrospective stress-forecasts of earthquakes and eruptions. Phys. Earth Planet. Inter. 245, 76–87. doi: 10.1016/j.pepi.2015.05.008

Crampin, S., Gao, Y., De Santis, A., 2013. A few earthquake conundrums resolved. J. Asian Earth Sci. 62, 501–509. doi: 10.1016/j.jseaes.2012.10.036

Crampin, S., Gao, Y., Peacock, S., 2008. Stress-forecasting (not predicting) earthquakes: A paradigm shift. Geology 36, 427. doi: 10.1130/G24643A.1

Crampin, S., Peacock, S., Gao, Y., Chastin, S., 2004. The scatter of time-delays in shear-wave splitting above small earthquakes. Geophys. J. Int. 156, 39–44. doi: 10.1111/j.1365-246X.2004.02040.x

Crampin, S., Volti, T., Stefánsson, R., 1999. A successfully stress-forecast earthquake. Geophys. J. Int. 138(*1*), F1-F5. doi: 10.1046/j.1365-246x.1999.00891.x

Crampin, S., Zatsepin, S., 1997. Modelling the compliance of crustal rock—II. Response to temporal changes before earthquakes. Geophys. J. Int. 129.

Crotwell, H.P., Owens, T.J., Ritsema, J., 1999. The TauP Toolkit: Flexible Seismic Travel-time and Ray-path Utilities. Seismol. Res. Lett. 70, 154–160. doi: 10.1785/gssrl.70.2.154

Das, R., Rai, S.S., 2017. Extensive seismic anisotropy in the lower crust of Archean metamorphic terrain, South India, inferred from ambient noise tomography. Tectonophysics 694, 164–180. doi: 10.1016/j.tecto.2016.12.002

De Martini, P.M., Pantosti, D., Palyvos, N., Lemeille, F., McNeill, L., Collier, R., 2004. Slip rates of the Aigion and Eliki Faults from uplifted marine terraces, Corinth Gulf, Greece. Comptes Rendus - Geosci. 336, 325–334. doi: 10.1016/j.crte.2003.12.006

Del Pezzo, E., Bianco, F., Petrosino, S., Saccorotti, G., 2004. Changes in the Coda Decay Rate and Shear-Wave Splitting Parameters Associated with Seismic Swarms at Mt . Vesuvius , Italy. Bull. Seismol. Soc. Am. 94, 439–452.

Diaz, J., Gallart, J., Villaseñor, A., Mancilla, F., Pazos, A., Córdoba, D., Pulgar, J.A., Ibarra, P., Harnafi, M., 2010. Mantle dynamics beneath the Gibraltar Arc (western Mediterranean) from shear-wave splitting measurements on a dense seismic array. Geophys. Res. Lett. 37, 1–5. doi: 10.1029/2010GL044201

Ding, Y., Peng, Y., Li, J., 2018. Cluster Analysis of Earthquake Ground-Motion Records and Characteristic Period of Seismic Response Spectrum. J. Earthq. Eng. 1–22. doi: 10.1080/13632469.2018.1453420

Doutsos, T., Poulimenos, G., 1992. Geometry and kinematics of active faults and their seismotectonic significance in the western Corinth-Patras rift (Greece). J. Struct. Geol. 14, 689–699. doi: 10.1016/0191-8141(92)90126-H

Drakatos, G., Karastathis, V., Makris, J., Papoulia, J., Stavrakakis, G., 2005. 3D crustal structure in the neotectonic basin of the Gulf of Saronikos (Greece). Tectonophysics 400, 55–65. doi: 10.1016/j.tecto.2005.02.004

Dresen, G., Guéguen, Y., 2004. Chapter 4 - Damage and Rock Physical Properties, in: Guéguen, Y., Boutéca, M.B.T.-I.G. (Eds.), Mechanics of Fluid-Saturated Rocks. Academic Press, pp. 169–217. doi: 10.1016/S0074-6142(03)80020-7

Druitt, T.H., 2014. New insights into the initiation and venting of the Bronze-Age eruption of Santorini (Greece), from component analysis. Bull. Volcanol. 76, 1–21. doi: 10.1007/s00445-014-0794-x

Duda, R.O., Hart, P.E., 1973. Pattern Classification and Scene Analysis, 1st ed. Wiley, New York.

Duverger, C., Godano, M., Bernard, P., Lyon-Caen, H., Lambotte, S., 2015. The 2003-2004 seismic swarm in the western Corinth rift: Evidence for a multiscale pore pressure diffusion process along a permeable fault system. Geophys. Res. Lett. 42, 7374–7382. doi: 10.1002/2015GL065298

Ekström, G., Nettles, M., Dziewoński, A.M., 2012. The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. Phys. Earth Planet. Inter. 200–201, 1–9. doi: 10.1016/J.PEPI.2012.04.002

Ellsworth, W.L., Llenos, A.L., McGarr, A.F., Michael, A.J., Rubinstein, J.L., Mueller, C.S., Petersen, M.D., Calais, E., 2015. Increasing seismicity in the U. S. midcontinent: Implications for earthquake hazard. Lead. Edge 34, 618–626. doi: 10.1190/tle34060618.1

Evangelidis, C.P., 2017. Seismic anisotropy in the Hellenic subduction zone: Effects of slab segmentation and subslab mantle flow. Earth Planet. Sci. Lett. 480, 97–106. doi: 10.1016/j.epsl.2017.10.003

Evangelidis, C.P., Liang, W.T., Melis, N.S., Konstantinou, K.I., 2011. Shear wave anisotropy beneath the Aegean inferred from SKS splitting observations. J. Geophys. Res. Solid Earth 116, 1–14. doi: 10.1029/2010JB007884

Evans, R., 1984. Effects of the free surface on shear wavetrains. Geophys. J. Int. 76, 165–172. doi: 10.1111/j.1365-246X.1984.tb05032.x

Everitt, B.S., Landau, S., Leese, M., Stahl, D., 2011. Cluster Analysis, 5th ed. Wiley. doi: 10.1002/9780470977811

Foumelis, M., Trasatti, E., Papageorgiou, E., Stramondo, S., Parcharidis, I., 2013. Monitoring santorini volcano (Greece) breathing from space. Geophys. J. Int. 193, 161–170. doi: 10.1093/gji/ggs135

Fukao, Y., 1984. Evidence from core-reflected shear waves for anisotropy in the Earth's mantle. Nature. doi: 10.1038/309695a0

Ganas, A., Chousianitis, K., Batsi, E., Kolligri, M., Agalos, A., Chouliaras, G., Makropoulos, K., 2013b. The January 2010 Efpalion earthquakes (Gulf of Corinth, Central Greece): Earthquake interactions and blind normal faulting. J. Seismol. 17, 465–484. doi: 10.1007/s10950-012-9331-6

Ganas, A., Oikonomou, A.I., Tsimi, C., 2013a. NOAFAULTS : a Digital Database for Active Faults in Greece. Bull. Geol. Soc. Greece XLVII, 518–530.

Gao, Y., Crampin, S., 2006. A stress-forecast earthquake (with hindsight), where migration of source earthquakes causes anomalies in shear-wave polarisations. Tectonophysics 426, 253–262. doi: 10.1016/j.tecto.2006.07.013

Gao, Y., Crampin, S., 2004. Observations of stress relaxation before earthquakes. Geophys. J. Int. 157, 578–582. doi: 10.1111/j.1365-246X.2004.02207.x

Giannopoulos, D., Sokos, E., Konstantinou, K.I., Tselentis, G. - A., 2015. Shear wave splitting and VP/VS variations before and after the Efpalio earthquake sequence, western Gulf of Corinth, Greece. Geophys. J. Int. 200, 1436–1448. doi: 10.1093/gji/ggu467

Godano, M., Deschamps, A., Lambotte, S., Lyon-Caen, H., Bernard, P., Pacchiani, F., 2014. Focal mechanisms of earthquake multiplets in the western part of the Corinth Rift (Greece): influence of the velocity model and constraints on the geometry of the active faults. Geophys. J. Int. 197(*3*), 1660-1680. doi: 10.1093/gji/ggu059

Guichard, F., Carey, S., Arthur, M.A., Sigurdsson, H., Arnold, M., 1993. Tephra from the Minoan eruption of Santorini in sediments of the Black Sea. Nature 363, 610.

Hatzfeld, D., Karakostas, V., Ziazia, M., Kassaras, I., Papadimitriou, E., Makropoulos, K., Voulgaris, N., Papaioannou, C., 2000. Microseismicity and faulting geometry in the Gulf of Corinth (Greece). Geophys. J. Int. 141, 438–456. doi: 10.1046/j.1365-246x.2000.00092.x

Helbig, K., 1984. Anisotropy and dispersion in periodically layered media. Geophysics 49, 364–373. doi: 10.1190/1.1441672

Helffrich, G., Wookey, J., Bastow, I., 2013. The Seismic Analysis Code: A Primer and User's Guide. Cambridge University Press, Cambridge. doi: 10.1017/CBO9781139547260

Houghton, S.L., 2003. New ^{234}U-^{230}Th coral dates from the western Gulf of Corinth: Implications for extensional tectonics. Geophys. Res. Lett. 30, 2013. doi: 10.1029/2003GL018112

Houser, C., Masters, G., Shearer, P., Laske, G., 2008. Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms. Geophys. J. Int. 174, 195–212. doi: 10.1111/j.1365-246X.2008.03763.x

Hunter, J.D., 2007. Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95. doi: 10.1109/MCSE.2007.55

Jammalamadaka, S.R., SenGupta, A., 2001. Topics in Circular Statistics, Series on Multivariate Analysis. WORLD SCIENTIFIC. doi: doi:10.1142/4031

Johnson, J.H., Savage, M.K., 2012. Tracking volcanic and geothermal activity in the Tongariro Volcanic Centre, New Zealand, with shear wave splitting tomography. J. Volcanol. Geotherm. Res. 223–224, 1–10. doi: 10.1016/j.jvolgeores.2012.01.017

Johnson, J.H., Savage, M.K., Townend, J., 2011. Distinguishing between stress-induced and structural anisotropy at Mount Ruapehu volcano, New Zealand. J. Geophys. Res. Solid Earth 116, 1–18. doi: 10.1029/2011JB008308

Kapetanidis, V., 2017. Spatiotemporal Patterns of Microseismicity for the Identification of Active Fault Structures using Seismic Waveform Cross-Correlation and Double-Difference Relocation. PhD Thesis, National and Kapodistrian University of Athens.

Kapetanidis, V., Deschamps, A., Papadimitriou, P., Matrullo, E., Karakonstantis, A., Bozionelos, G., Kaviris, G., Serpetsidaki, A., Lyon-Caen, H., Voulgaris, N., Bernard, P., Sokos, E., Makropoulos, K., 2015. The 2013 earthquake swarm in Helike, Greece: seismic activity at the root of old normal faults. Geophys. J. Int. 202, 2044–2073. doi: 10.1093/gji/ggv249

Karakonstantis, A., 2017. Simulation of the 3D structure of the Crust and Upper Mantle in Greece using Seismic Tomography. PhD Thesis, National and Kapodistrian University of Athens (in Greek)

Karakostas, V., Karagianni, E., Paradisopoulou, P., 2012. Space-time analysis, faulting and triggering of the 2010 earthquake doublet in western Corinth Gulf. Nat. Hazards 63, 1181–1202. doi: 10.1007/s11069-012-0219-0

Kassaras, I., Kapetanidis, V., 2018. Resolving the Tectonic Stress by the Inversion of Earthquake Focal Mechanisms. Application in the Region of Greece. A Tutorial, in: D'Amico, S. (Ed.), Moment Tensor Solution. Springer, Cham, pp. 405–452. doi: 10.1007/978-3-319-77359-9_19

Kaviris, G., 2003. Study of seismic source properties of the eastern Gulf of Corinth. PhD Thesis, National and Kapodistrian University of Athens (in Greek).

Kaviris, G., Fountoulakis, I., Spingos, I., Millas, C., Papadimitriou, P., 2018c. Mantle dynamics beneath Greece from SKS and PKS seismic anisotropy study. Acta Geophys. 66(*6*), 1341-1357. doi: 10.1007/s11600-018-0225-z

Kaviris, G., Millas, C., Spingos, I., Kapetanidis, V., Fountoulakis, I., Makropoulos, K., 2018d. Massive shear-wave splitting measurements towards an upper crust seismic anisotropy study in the Western Gulf of Corinth (Greece), in: 36th General Assembly of the European Seismological Comission, Books of Abstracts. p. 150.

Kaviris, G., Millas, C., Spingos, I., Kapetanidis, V., Fountoulakis, I., Papadimitriou, P., Voulgaris, N., Makropoulos, K., 2018a. Observations of shear-wave splitting parameters in the Western Gulf of Corinth focusing on the 2014 M_w = 5.0 earthquake. Phys. Earth Planet. Inter. 282, 60–76. doi: 10.1016/j.pepi.2018.07.005

Kaviris, G., Papadimitriou, P., Kravvariti, P., Kapetanidis, V., Karakonstantis, A., Voulgaris, N., Makropoulos, K., 2015. A detailed seismic anisotropy study during the 2011-2012 unrest period in the Santorini Volcanic Complex. Phys. Earth Planet. Inter. 238, 51–88. doi: 10.1016/j.pepi.2014.11.002

Kaviris, G., Papadimitriou, P., Makropoulos, K., 2008. An Overview of Anisotropy Studies in Central Greece using recordings around the Gulf of Corinth (Greece) and aftershocks of the 1999 Athens Earthquake, in: 31st General Assembly of the European Seismological Commission. Hersonissos, Crete, pp. 215–223.

Kaviris, G., Spingos, I., Kapetanidis, V., Papadimitriou, P., Voulgaris, N., Makropoulos, K., 2017. Upper crust seismic anisotropy study and temporal variations of shear-wave splitting parameters in the Western Gulf of Corinth (Greece) during 2013. Phys. Earth Planet. Inter. 269, 148–164. doi: doi: 10.1016/j.pepi.2017.06.006

Kaviris, G., Spingos, I., Millas, C., Kapetanidis, V., Fountoulakis, I., Papadimitriou, P., Voulgaris, N., Drakatos, G., 2018b. Effects of the January 2018 seismic sequence on shear-wave splitting in the upper crust of Marathon (NE Attica, Greece). Phys. Earth Planet. Inter. 285, 45–58. doi: 10.1016/j.pepi.2018.10.007

Keranen, K.M., Savage, H.M., Abers, G.A., Cochran, E.S., 2013. Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 M_w 5.7 earthquake sequence. Geology 41, 699–702. doi: 10.1130/G34045.1

Klein, F.W., 2002. User's Guide to HYPOINVERSE-2000, a Fortran Program to Solve for Earthquake Locations and Magnitudes. U.S. Geol. Surv. Open File Rep. 02-171 123 pp.

Koukouvelas, I.K., 1998. The Egion Fault, earthquake-related and long-term deformation, Gulf of Corinth, Greece. J. Geodyn. 26, 501–513. doi: 10.1016/S0264-3707(97)00046-X

Konstantinou, K.I., Evangelidis, C.P., Liang, W., Melis, N.S., Kalogeras, I., 2013. Seismicity , V_p/V_s and shear wave anisotropy variations during the 2011 unrest at Santorini caldera , southern Aegean. J. Volcanol. Geotherm. Res. 267, 57–67. doi: 10.1016/j.jvolgeores.2013.10.001

Koukouvelas, I.K., Doutsos, T.T., 1996. Implications of structural segmentation during earthquakes: The 1995 Egion earthquake, Gulf of Corinth, Greece. J. Struct. Geol. 18, 1381–1388. doi: 10.1016/S0191-8141(96)00071-5

Koukouvelas, I.K., Katsonopoulou, D., Soter, S., Xypolias, P., 2005. Slip rates on the Helike Fault, Gulf of Corinth, Greece: new evidence from geoarchaeology. Terra Nov. 17, 158–164. doi: 10.1111/j.1365-3121.2005.00603.x

Kouskouna, V., Sakkas, G., 2013. The University of Athens Hellenic Macroseismic Database (HMDB.UoA): Historical earthquakes. J. Seismol. 17, 1253–1280. doi: 10.1007/s10950-013-9390-3

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C., Wassermann, J., 2015. ObsPy: a bridge for seismology into the scientific Python ecosystem. Comput. Sci. Discov. 8. doi: 10.1088/1749-4699/8/1/014003

Lagios, E., Sakkas, V., Novali, F., Bellotti, F., Ferretti, A., Vlachou, K., Dietrich, V., 2013. SqueeSAR™ and GPS ground deformation monitoring of Santorini Volcano (1992-2012): Tectonic implications. Tectonophysics 594, 38–59. doi: 10.1016/j.tecto.2013.03.012

Lai, H.L., Gibson, R.L., Lee, K.J., 2009. Quasi-shear wave ray tracing by wavefront construction in 3-D, anisotropic media. J. Appl. Geophys. 69, 82–95. doi: 10.1016/j.jappgeo.2009.06.002

Latorre, D., Virieux, J., Monfret, T., Monteiller, V., Vanorio, T., Got, J.L., Lyon-Caen, H., 2004. A new seismic tomography of Aigion area (Gulf of Corinth, Greece) from the 1991 data set. Geophys. J. Int. 159, 1013–1031. doi: 10.1111/j.1365-246X.2004.02412.x

Lei, J., 2005. A method for non-orthogonal seismic polarization-vector separation. Geophys. J. Int. 162, 965–974. doi: 10.1111/j.1365-246X.2005.02709.x

Li, X.Y., MacBeth, C., Crampin, S., 1998. Interpreting non-orthogonal split shear waves for seismic anisotropy in multicomponent VSPS. Geophys. Prospect. 46, 1–27. doi: 10.1046/j.1365-2478.1998.750311.x

Lyon-Caen, H., Papadimitriou, P., Deschamps, A., Bernard, P., Makropoulos, K., Pacchiani, F., Patau, G., 2004. First results of the CRLN seismic network in the western Corinth Rift: evidence for old-fault reactivation. Comptes Rendus - Geosci. 336, 343–351. doi: 10.1016/j.crte.2003.12.004

Marinatos, S., 1939. The Volcanic Destruction of Minoan Crete. Antiquity 13, 425–439. doi: 10.1017/S0003598X00028088

Makropoulos, K., Kaviris, G., Kouskouna, V., 2012. An updated and extended earthquake catalogue for Greece and adjacent areas since 1900. Nat. Hazards Earth Syst. Sci. 12, 1425–1430. doi: 10.5194/nhess-12-1425-2012

McNeill, L.C., Cotterill, C.J., Henstock, T.J., Bull, J.M., Stefatos, T.J., Collier, R.E.L., Papatheoderou, G., Ferentinos, G., Hicks, S.E., 2005. Active faulting within the offshore western Gulf of Corinth, Greece: Implications for models of continental rift deformation. Geology 33, 241–244. doi: 10.1130/G21127.1

Mesimeri, M., Karakostas, V., Papadimitriou, E., Schaff, D., Tsaklidis, G., 2016. Spatio-temporal properties and evolution of the 2013 Aigion earthquake swarm (Corinth Gulf, Greece). J. Seismol. 20, 595–614. doi: 10.1007/s10950-015-9546-4

Mesimeri, M., Karakostas, V., Papadimitriou, E., Tsaklidis, G., Jacobs, K., 2018. Relocation of recent seismicity and seismotectonic properties in the Gulf of Corinth (Greece). Geophys. J. Int. 212, 1123–1142. doi: 10.1093/gji/ggx450

Milligan, G.W., Cooper, M.C., 1988. A study of standardization of variables in cluster analysis. J. Classif. 5, 181–204. doi: 10.1007/BF01897163

Moretti, I., Sakellariou, D., Lykousis, V., Micarelli, L., 2003. The Gulf of Corinth: An active half graben? J. Geodyn. 36, 323–340. doi: 10.1016/S0264-3707(03)00053-X

Nolte, K.A., Tsoflias, G.P., Bidgoli, T.S., Watney, W.L., 2017. Shear-wave anisotropy reveals pore fluid pressure–induced seismicity in the U.S. midcontinent. Sci. Adv. 3, e1700443. doi: 10.1126/sciadv.1700443

Nomikou, P., Hübscher, C., Papanikolaou, D., Farangitakis, P.G., Ruhnau, M., Lampridou, D., 2017. Expanding extension, subsidence and lateral segmentation within the Santorini - Amorgos basins during Quaternary: Implications for the 1956 Amorgos events, central - south Aegean Sea, Greece. Tectonophysics 722, 138–153. doi: 10.1016/j.tecto.2017.10.016

Nomikou, P., Parks, M.M., Papanikolaou, D., Pyle, D.M., Mather, T.A., Carey, S., Watts, A.B., Paulatto, M., Kalnins, M.L., Livanos, I., Bejelou, K., Simou, E., Perros, I., 2014. The emergence and growth of a submarine volcano: The Kameni islands, Santorini (Greece). GeoResJ 1–2, 8–18. doi: 10.1016/j.grj.2014.02.002

Oliphant, T.E., 2007. Python for Scientific Computing. Comput. Sci. Eng. 9, 10–20. doi: 10.1109/MCSE.2007.58

Pantosti, D., De Martini, P.M., Koukouvelas, I., Stamatopoulos, L., Palyvos, N., Pucci, S., Lemeille, F., Pavlides, S., 2004. Palaeoseismological investigations of the Aigion Fault (Gulf of Corinth, Greece). Comptes Rendus Geosci. 336, 335–342. doi: 10.1016/j.crte.2003.12.005

Papadimitriou, P., Kapetanidis, V., Karakonstantis, A., Kaviris, G., Voulgaris, N., Makropoulos, K., 2015. The Santorini Volcanic Complex: A detailed multi-parameter seismological approach with emphasis on the 2011-2012 unrest period. J. Geodyn. 85, 32–57. doi: 10.1016/j.jog.2014.12.004

Papadimitriou, P., Kaviris, G., Makropoulos, K., 1999. Evidence of shear-wave splitting in the eastern Corinthian Gulf (Greece). Phys. Earth Planet. Inter. 114, 3–13.

Papanikolaou, D.J., Royden, L.H., 2007. Disruption of the Hellenic arc: Late Miocene extensional detachment faults and steep Pliocene-Quaternary normal faults-Or what happened at Corinth? Tectonics 26, TC5003. doi: 10.1029/2006TC002007

Papazachos, V., Papazachou, K., 2003. The earthquakes of Greece. Ziti Publications Co (in Greek).

Pareschi, M.T., Favalli, M., Boschi, E., 2006. Impact of the Minoan tsunami of Santorini: Simulated scenarios in the Eastern Mediterranean. Geophys. Res. Lett. 33, 1–6. doi: 10.1029/2006GL027205

Pedregosa, F., Varoqaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blonddel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830. doi: 10.1007/s13398-014-0173-7.2

Peng, Z., Ben-Zion, Y., 2004. Systematic analysis of crustal anisotropy along the Karadere-Düzce branch of the North Anatolian fault. Geophys. J. Int. 159, 253–274. doi: 10.1111/j.1365-246X.2004.02379.x

Pérez, F., Granger, B.E., 2007. IPython: A System for Interactive Scientific Computing. Comput. Sci. Eng. 9, 21–29. doi: 10.1109/MCSE.2007.53

Perissoratis, C., 1995. The Santorini volcanic complex and its relation to the stratigraphy and structure of the Aegean arc, Greece. Mar. Geol. 128, 37–58. doi: 10.1016/0025-3227(95)00090-L

Piccinini, D., Pastori, M., Margheriti, L., 2013. ANISOMAT+: An automatic tool to retrieve seismic anisotropy from local earthquakes. Comput. Geosci. 56, 62–68. doi: 10.1016/j.cageo.2013.01.012

Plesinger, A., Hellweg, M., Seidl, D., 1986. Interactive high-resolution polarization analysis of broad-band seismograms. J. Geophys. 59, 129–139.

Potanina, M.G., Smirnov, V.B., Bernard, P., 2011. Patterns of seismic swarm activity in the Corinth Rift in 2000–2005. Izv. Phys. Solid Earth 47, 610–622. doi: 10.1134/S106935131106005X

Pyle, D.M., 1997. The global impact of the minoan eruption of Santorini, Greece. Environ. Geol. 30, 59–61. doi: 10.1007/s002540050132

Reiss, M.C., Rümpker, G., 2017. SplitRacer: MATLAB Code and GUI for Semiautomated Analysis and Interpretation of Teleseismic Shear-Wave Splitting. Seismol. Res. Lett. 88, 392–409. doi: 10.1785/0220160191

Reuter, H.I., Nelson, A., Jarvis, A., 2007. An evaluation of void-filling interpolation methods for SRTM data. Int. J. Geogr. Inf. Sci. 21, 983–1008. doi: 10.1080/13658810601169899

Rigo, A., Lyon-Caen, H., Armijo, R., Deschamps, A., Hatzfeld, D., Makropoulos, K., Papadimitriou, P., Kassaras, I., 1996. A microseismic study in the western part of the Gulf of Corinth (Greece): implications for large-scale normal faulting mechanisms. Geophys. J. Int. 126, 663–688.

Riverbank Computing, 2018. PyQt5. www.riverbankcomputing.com/software/pyqt/intro

Sadri, M., Riahi, M.A., 2010. Ray tracing and amplitude calculation in anisotropic layered media. Geophys. J. Int. 180, 1170–1180. doi: 10.1111/j.1365-246X.2009.04464.x

Salat, D.H., 2014. Chapter 12 - Diffusion Tensor Imaging in the Study of Aging and Age-Associated Neural Disease, in: Johansen-Berg, H., Behrens, T.E.J.B.T.-D.M.R.I. (Second E. (Eds.), . Academic Press, San Diego, pp. 257–281. doi: doi: 10.1016/B978-0-12-396460-1.00012-3

Sandanayake, N., Bale, R., 2011. Application of Shear-wave Splitting Analysis to Fracture Characterization for a Shaunavon Tight Oil Reservoir. GeoConvention 1–4.

Savage, M.K., Silver, P.G., 1993. Mantle deformation and tectonics: constraints from seismic anisotropy in the western United States. Phys. Earth Planet. Inter. 78, 207–227. doi: 10.1016/0031-9201(93)90156-4

Savage, M.K., Wessel, A., Teanby, N. A., Hurst, A. W., 2010. Automatic measurement of shear wave splitting and applications to time varying anisotropy at Mount Ruapehu volcano, New Zealand. J. Geophys. Res. Solid Earth 115. doi: 10.1029/2010JB007722

Schorlemmer, D., Euchner, F., Kästli, P., Saul, J., 2011. QuakeML: status of the XML-based seismological data exchange format. Ann. Geophys. 54, 59–65. doi: 10.4401/ag-4874

Sharma, A., Baruah, S., Piccinini, D., Saikia, S., Phukan, M.K., Chetia, M., Kayal, J.R., 2017. Crustal seismic anisotropy beneath Shillong plateau - Assam valley in North East India: Shear-wave splitting analysis using local earthquakes. Tectonophysics 717, 425–432. doi: 10.1016/j.tecto.2017.08.027

Shearer, P., Hauksson, E., Lin, G., 2005. Southern California hypocenter relocation with waveform cross-correlation, part 2: Results using source-specific station terms and cluster analysis. Bull. Seismol. Soc. Am. 95, 904–915. doi: 10.1785/0120040168

Shearer, P.M., 2009. Introduction to Seismology, 2nd ed. Cambridge University Press.

Sibson, R.H., 1994. Crustal stress, faulting and fluid flow. Geol. Soc. London, Spec. Publ. 78, 69–84. doi: 10.1144/GSL.SP.1994.078.01.07

Silver, P.G., Chan, W.W., 1991. Shear Wave Splitting and Subcontinental Mantle Deformation. J. Geophys. Res. 96(*B10*), 16429–16454. doi: 10.1029/91JB00899

Skourtsos, E., Kranis, H., 2009. Structure and evolution of the western Corinth Rift, through new field data from the Northern Peloponnesus. Geol. Soc. London, Spec. Publ. 321, 119–138. doi: 10.1144/SP321.6

Smoot, G.F., Gorenstein, M. V., Muller, R.A., 1977. Detection of Anisotropy in the Cosmic Blackbody Radiation. Phys. Rev. Lett. 39, 898–901. doi: 10.1103/PhysRevLett.39.898

Sokos, E., Zahradník, J., Kiratzi, A., Janský, J., Gallovič, F., Novotny, O., Kostelecký, J., Serpetsidaki, A., Tselentis, G. -A., 2012. The January 2010 Efpalio earthquake sequence in the western Corinth Gulf (Greece). Tectonophysics 530–531, 299–309. doi: 10.1016/j.tecto.2012.01.005

Soter, S., Katsonopoulou, D., 2011. Submergence and uplift of settlements in the area of Helike, Greece, from the Early Bronze Age to late antiquity. Geoarchaeology 26, 584–610. doi: 10.1002/gea.20366

Spingos, I., Kaviris, G., Kassaras, I., 2018. Observations of ambient noise induced by Hurricane Katrina in the southern USA, in: 36th General Assembly of the European Seismological Comission, Books of Abstracts. Valetta, Malta, p. 405.

Stanley, D.J., Sheng, H., 1986. Volcanic shards from Santorini (Upper Minoan ash) in the Nile Delta, Egypt. Nature 320, 733.

Stephens, C.D., Page, R.A., Lahr, J.C., 1990. Reflected and mode-converted seismic waves within the shallow aleutian subduction zone, southern Kenai Peninsula, Alaska. J. Geophys. Res. 95, 6883–6897. doi: 10.1029/JB095iB05p06883

Sun, C.-T., Jin, Z.-H., 2012. Fracture Mechanics, 1st ed. Academic Press. doi: 10.1016/C2009-0-63512-1

Syuhada, S., Hananto, N.D., Abdullah, C.I., Puspito, N.T., Yudistira, T., Anggono, T., 2017. Study on 2-D Crustal Shear Wave Splitting Tomography along The Sunda-Banda Arc Transition Zone. IOP Conf. Ser. Earth Environ. Sci. 62, 012054. doi: 10.1088/1755-1315/62/1/012054

Teanby, N., Kendall, J.-M., van der Baan, M., 2004. Automation of shear-wave splitting measurements using cluster analysis. Bulliten Seismol. Soc. Am. 94, 453–463. doi: 10.1785/0120030123

The Mathworks Inc., 2018. MATLAB. www.mathworks.com

The Qt Company, 2018. Qt. www.qt.io

Tiberi, C., Lyon-Caen, H., Hatzfeld, D., Achauer, U., Karagianni, E., Kiratzi, A., Louvari, E., Panagiotopoulos, D., Kassaras, I., Kaviris, G., Makropoulos, K., Papadimitriou, P., 2000. Crustal and upper mantle structure beneath the Corinth rift (Greece) from a teleseismic tomography study. J. Geophys. Res. 105(*B12*), 28159-28171. doi: 10.1029/2000JB900216

Tselentis, G.A., Makropoulos, K.C., 1986. Rates of crustal deformation in the Gulf of Corinth (central Greece) as determined from seismicity. Tectonophysics 124(*1-2*), 55-57. doi: 10.1016/0040-1951(86)90137-X

Tsimi, C., Ganas, A., Soulakellis, N., Kairis, O., Valmis, S., 2007. Morphotectonics of the Psathopyrgos active fault , western Corinth Rift , central Greece . Bull. Geol. Soc. Greece vol. XXXVII, 2007 Proc. 11th Int. Congr. Athens, May, 2007 XXXVII.

Valcke, S.L.A., Casey, M., Lloyd, G.E., Kendall, J.-M., Fisher, Q.J., 2006. Lattice preferred orientation and seismic anisotropy in sedimentary rocks. Geophys. J. Int. 166, 652–666. doi: 10.1111/j.1365-246X.2006.02987.x

Vallianatos, F., Michas, G., Papadakis, G., Tzanis, A., 2013. Evidence of non-extensivity in the seismicity observed during the 2011-2012 unrest at the Santorini volcanic complex, Greece. Nat. Hazards Earth Syst. Sci. 13, 177–185. doi: 10.5194/nhess-13-177-2013

van Rossum, G., 1995. Python tutorial, Technical Report CS-R9526. Amsterdam.

Vespa, M., Keller, J., Gertisser, R., 2006. Interplinian explosive activity of Santorini volcano (Greece) during the past 150,000 years. J. Volcanol. Geotherm. Res. 153, 262–286. doi: 10.1016/j.jvolgeores.2005.12.009

Vinnik, L.P., Farra, V., Romanowicz, B., 1989. Azimuthal Anisotropy in the Earth From Observations of SKS At Geoscope and Nars Broadband Stations. Bull. Seismol. Soc. Am. 79, 1542–1558.

Voulgaris, N., 1991. Investigation of crustal structure in Western Greece (Zakynthos - Western Peloponnese). PhD Thesis, National and Kapodistrian University of Athens (in Greek).

Waldhauser, F., 2001. hypoDD -- A Program to Compute Double-Difference Hypocenter Locations. USGS Open File Rep. 01-113.

Walpole, J., Wookey, J., Masters, G., Kendall, J.M., 2014. A uniformly processed data set of SKS shear wave splitting measurements: A global investigation of upper mantle anisotropy beneath seismic stations. Geochemistry, Geophys. Geosystems 15, 1991–2010. doi: 10.1002/2014GC005278

Walsh, E., Arnold, R., Savage, M.K., 2013. Silver and Chan revisited. J. Geophys. Res. Solid Earth 118, 5500–5515. doi: 10.1002/jgrb.50386

Wang, Y., 2014. Seismic ray tracing in anisotropic media: A modified Newton algorithm for solving highly nonlinear systems. Geophysics 79, T1–T7. doi: 10.1190/geo2013-0110.1

Lynn Watney, W., 2014. Wellington, Kansas CO_2 Sequestration Monitoring. International Federation of Digital Seismograph Networks. Other/Seismic Network. doi: 10.7914/SN/ZA_2014

Wessel, P., Smith, W.H.F., 1991. Free software helps map and display data. Eos, Trans. Am. Geophys. Union 72, 441–446. doi: 10.1029/90EO00319

Wüstefeld, A., Al-Harrasi, O., Verdon, J.P., Wookey, J., Kendall, J.M., 2010. A strategy for automated analysis of passive microseismic data to image seismic anisotropy and fracture characteristics. Geophys. Prospect. 58, 755–773. doi: 10.1111/j.1365-2478.2010.00891.x

Wüstefeld, A., Bokelmann, G., Zaroli, C., Barruol, G., 2008. SplitLab: A shear-wave splitting environment in Matlab. Comput. Geosci. 34, 515–528. doi: 10.1016/j.cageo.2007.08.002

Yardley, G.S., Crampin, S., 1991. Extensive-dilatancy anisotropy: relative information in VSPs and reflection surveys. Geophys. Prospect. 39, 337–355. doi: 10.1111/j.1365-2478.1991.tb00316.

Zatsepin, S., Crampin, S., 1997. Modelling the compliance of crustal rock—I. Response of shear-wave splitting to differential stress. Geophys. J. Int. 129, 477–494.

Zhang, H., Liu, Y., Thurber, C., Roecker, S., 2007. Three-dimensional shear-wave splitting tomography in the Parkfield, California, region. Geophys. Res. Lett. 34, 1–6. doi: 10.1029/2007GL031951