

Toward a Three-Dimensional Crustal Structure of the Dead Sea region from Local Earthquake Tomography

Thesis submitted to the Senate of Tel-Aviv University
for the degree of Doctor of Philosophy

by
Freddy Aldersons

Under the supervision of
Prof. Zvi Ben-Avraham (TAU)
and Prof. Eduard Kissling (ETH)

June 2004

To Marilena

Contents

Abstract		v
Chapter 1	Introduction	1
Chapter 2	Lower-crustal strength under the Dead Sea basin from local earthquake data and rheological modeling	4
	2.1 Introduction	4
	2.2 Geological setting	6
	2.3 Seismicity	7
	2.4 Rheology	15
	2.5 Discussion and conclusions	19
	Acknowledgments	21
Chapter 3	Methods of a new picking system: MannekenPix	22
	3.1 Automatic picking	23
	3.2 MannekenPix	25
	3.3 The four steps of MannekenPix	27
	3.3.1 Wiener filter	28
	Linear least-squares optimal filtering of stationary time series	28
	Stochastic and data-adaptive Wiener filters	31
	Power spectrum estimation	33
	Classical methods of power spectrum estimation	34
	The Maximum Entropy Method (MEM)	38
	Burg's algorithm	40
	Selection of the model order in MEM	42
	Properties of an MEM spectrum	43
	Power spectrum estimation in MannekenPix	47
	Application of the Wiener filter	48
	Wiener filter and deconvolution applied to a synthetic example	49
	3.3.2 Automatic picking	54
	3.3.3 Delay corrections	56
	3.3.4 Weighting	61
	Discriminant analysis	62
	The discriminating variables of MannekenPix	63

Chapter 4	High-quality automatic picking of local earthquake data from the Dead Sea region	67
	4.1 Methodology	68
	4.2 Data set characterization	71
	4.3 Picking rate and accuracy	76
	4.4 Picking uncertainty	82
	4.5 A simple relocation	89
Chapter 5	Conclusions	91
	Lower-crustal strength under the Dead Sea basin	91
	A new picking system: MannekenPix	92
	Deconvolution	93
	Application of MannekenPix to the Dead Sea region	94
	Application of MannekenPix to the seismicity of Italy	95
Appendix A	Abbreviations	97
Appendix B	Source function parameters and seismograph response of the Wiener-filtered synthetic example (Figure 3.6)	99
Appendix C	GII and JSO Station Coordinates	107
References		110
Acknowledgments		122

Abstract

We studied the local seismicity of the Dead Sea basin for the period 1984-1997. Sixty percent of well-constrained microearthquakes ($M_L \leq 3.2$) nucleated at depths of 20-32 km and more than 40 percent occurred below the depth of peak seismicity situated at 20 km. With the Moho at 32 km, the upper mantle appeared to be aseismic during the 14-year data period. A relocation procedure involving the simultaneous use of three regional velocity models reveals that the distribution of focal depths in the Dead Sea basin is stable. The lower-crustal seismicity under the basin is not an artifact created by strong lateral velocity variations or data-related problems. An upper bound depth uncertainty of ± 5 km is estimated below 20 km, but for most earthquakes depth mislocations should not exceed ± 2 km. A lithospheric strength profile has been calculated. Based on a surface heat flow of 40 mWm^{-2} and a quartz-depleted lower crust, a narrow brittle to ductile transition might occur in the crust around 380°C at a depth of 31 km. For the upper mantle, the brittle to ductile transition occurs in the model at 490°C and at 44 km depth. The absence of micro-seismicity in the upper mantle remains difficult to explain.

Within the framework of local earthquake tomography, a new automatic picking system called MannekenPix has been developed in order to collect a high-quality set of picking data from the Dead Sea region. In another work, this data set would then become the input of a high-resolution travel-time tomographic study of the crustal structure of the basin. In the first step of MannekenPix, the seismograms are filtered by a high-fidelity Wiener filter in order to increase the signal-to-noise ratio of the P-wavetrain before picking. The Wiener filter of MannekenPix uses the maximum entropy method to estimate power spectral densities, a method particularly well adapted to short data segments. The picking is performed in the second step by the robust and versatile Baer-Kradolfer (Baer and Kradolfer, 1987) picking engine. If a valid pick has been found, a variable delay correction is then applied in the third step in order to reduce the inherent delay of the Baer-Kradolfer algorithm arrivals. In the fourth step, the weighting engine provides a statistical estimate of the uncertainty. For each data set, the weighting engine has to be calibrated first by a multiple discriminant analysis performed on user-supplied reference picks and weights.

By using MannekenPix to pick the seismograms from well-constrained earthquakes of the Dead Sea region, 526 (99.1 %) out of 531 earthquakes are locatable by automatic picking. Out of 15,250 seismograms, 6,889 (45.2 %) were routinely picked and 7,089 (46.5 %) were

automatically picked. On the calibration data set, the average discrepancy between the routine picking and the reference picking is -0.037 sec, with a standard deviation of 0.121 sec. The average discrepancy of the automatic picking is $+0.007$ sec with a standard deviation of only 0.066 sec. For the complete data set, $3,058$ phases (43.1%) of the $7,089$ automatic P picks fall into predicted weight 1 (absolute uncertainty not greater than 40 msec), $1,173$ (16.5%) fall into class 2 (absolute uncertainty between 40 msec and 80 msec), $1,554$ (21.9%) fall into class 3 (absolute uncertainty between 80 msec and 140 msec) and $1,304$ (18.4%) fall into class 4 (absolute uncertainty greater than 140 msec). The results from a totally out-of-sample set of reference picks and weights confirm the stability of predicted weights on unseen data.

Di Stefano et al. (2004, to be submitted) present the application of MannekenPix to about $240,000$ seismograms from nearly $29,000$ local earthquakes routinely recorded by the Italian national seismic network during the period 1998-2001. MannekenPix was able to produce $103,131$ P picks (73% of the $139,500$ routine onsets) from $23,108$ events out of $28,900$ events. About $17,130$ phases (17%) of the $103,131$ fall into class 1 (absolute uncertainty not greater than 0.1 sec) and $15,429$ (15%) fall into class 2 (absolute uncertainty between 0.1 sec and 0.2 sec). Results show that MannekenPix arrival times for classes 1 and 2 combined, produce a distribution of residual times with a much smaller standard deviation than for the CSI bulletin. This means that a significant increase in the quality of the data set has been gained, coming at the price of a reduced quantity. The application of MannekenPix to the very large and inhomogeneous dataset of waveforms recorded in Italy from 1998 to 2001 yields nearly $33,000$ high-quality weighted picks and polarities. By using MannekenPix we achieved our goal to build a new dataset of P-wave arrival times and polarities belonging to user-defined classes 1 and 2 (the highest qualities) with associated error estimations. The analysis of relocation residuals and the time versus distance plots for these classes show the effectiveness of the picking system. Although the hit rate of MannekenPix 1.6.2 applied to this large and noisy dataset did not exceed 75% of the rate of an expert seismologist, the consistency of computed arrivals and polarities and their rapid estimation are very suitable to substitute bulletin data, solving the typical problem of extending consistency and quality to large datasets, and saving a great amount of time. We believe that the improved seismic dataset is suitable to refine the seismic tomography and to improve the knowledge of the local and regional stress fields in the Italian peninsula through the determination of a large number of focal mechanisms for well-located events, extending the analysis to small magnitudes and past events.

Chapter 1

Introduction

The Dead Sea Transform (Figure 1.1a) is an intracontinental plate boundary resulting from the late-Cenozoic breakup of the Arabo-African continent. This boundary extends over 1,000 km from the zone of sea floor spreading at the southern tip of the Sinai Peninsula to the Taurus-Zagros zone of convergence in Turkey (Freund, 1965). The Dead Sea basin (Figure 1.1b) is an active pull-apart (Quennell, 1958) located along the Dead Sea Transform. The amount of left-lateral motion along the Transform in the Dead Sea region is estimated at 105 km (Quennell, 1958; Freund et al., 1970). Motion along the Transform started around 20 Ma ago during Miocene times (Bartov et al., 1980). Although the Dead Sea basin originated at the beginning of the Transform motion, it did not develop into a topographic low before early Pliocene (Garfunkel, 1997). Pliocene sediments are mostly composed of evaporites, among which halite is the main constituent. Deformation of Pliocene salt has created numerous diapirs in the basin (Neev and Hall, 1979). During Pleistocene times, the basin subsidence accelerated (ten Brink and Ben-Avraham, 1989) and allowed the accumulation of several kilometers of lacustrine clastics, carbonates and evaporites. Today the basin is a morphotectonic depression over 130 km long and 7-18 km wide. It is a seismically active section (van Eck and Hofstetter, 1989, 1990) along the Dead Sea Transform, for which some 4,000 years of combined archaeological, historical and instrumental seismicities are documented (Ben-Menahem, 1991). Recent articles on the tectonics of the Dead Sea basin can be found in Cloetingh and Ben-Avraham (2002).

The Dead Sea basin and its surroundings have been monitored locally during nearly 20 years by permanent digital seismic stations. In addition, the region has been the focus of detailed geophysical studies during past decades. These studies include bathymetry (Neev and Hall, 1979), seismic reflection (Neev and Hall, 1979; ten Brink and Ben-Avraham, 1989), seismic

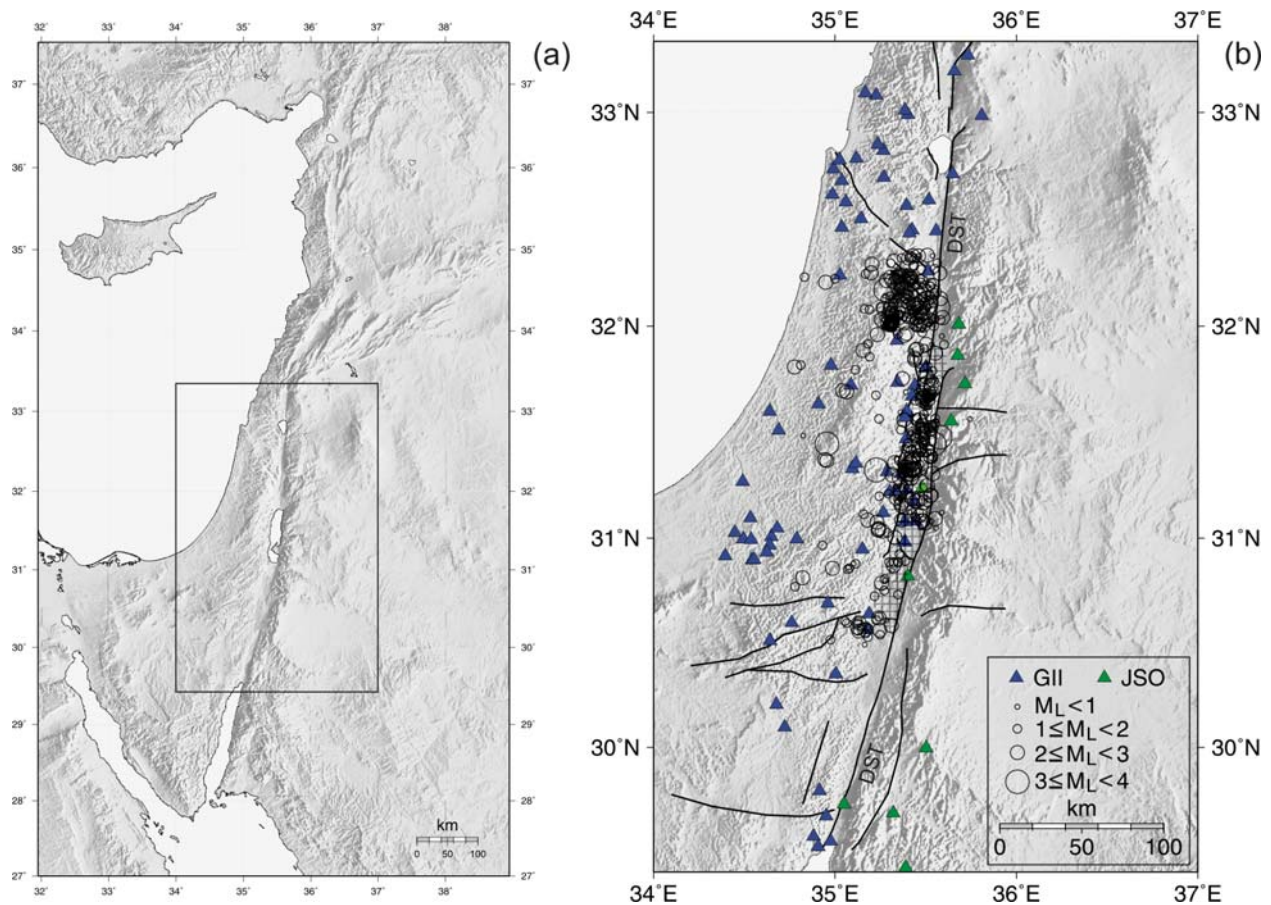


Figure 1.1 Regional setting of the Dead Sea basin. (a) The East-Mediterranean region. (b) The Dead Sea region. Five hundred thirty-one well-locatable earthquakes (1984-2001) in the vicinity of the Dead Sea basin, recorded by short-period stations (triangles) of GII (Israel) and JSO (Jordan). The square grid fill defines the Dead Sea basin. DST: Dead Sea Transform. Topographic shaded relief derived from Globe 1-km elevation data.

refraction (Ginzburg et al., 1981; Ginzburg and Ben-Avraham, 1997), wide-angle reflection-refraction (DESERT Group, 2004), gravity (ten Brink et al., 1990; ten Brink et al., 1993), magnetism (Frieslander and Ben-Avraham, 1989; Al-Zoubi and Ben-Avraham, 2002), heat flow (Ben-Avraham et al., 1978) and seismicity (van Eck and Hofstetter, 1989, 1990).

Local earthquake travel-time tomography can provide information in a depth range that cannot be imaged by seismic reflection methods, and it can resolve details of the crustal structure that are not accessible to regional studies. It could be a particularly helpful technique to image the lower crust under the Dead Sea basin, for which little is known and whose role in the development of the basin is poorly understood. Two local earthquake tomographic studies (Rabinowitz et al., 1996; Rabinowitz and Mart, 2000) have already been performed in the Dead Sea basin but with very small data sets.

The main purpose of this work is to collect a high-quality set of picking data from the Dead Sea region through the development of an automatic picking system. In another work, this data set would then become the input of a high-resolution travel-time tomographic study of the crustal structure of the basin and vicinity.

Routine earthquake locations in the Dead Sea region suggest that an anomalously high number of micro-earthquakes might have nucleated in the lower-crust under the basin. We studied the local seismicity of the Dead Sea basin for the period 1984-1997. This seismological study, complemented by a rheological strength profile of the lithosphere, is presented in Chapter 2.

The methods underlying the new picking system called MannekenPix are described in Chapter 3. Built around the Baer-Kradolfer (1987) single-trace picking algorithm, MannekenPix first reduces the noise on the seismograms by applying an adaptive high-fidelity Wiener filter where power spectral densities are estimated from short data segments. The Wiener filter of MannekenPix uses the maximum entropy method (MEM) which is particularly well adapted to short data segments. The Baer-Kradolfer (1987) picking algorithm is described next and the delay it introduces is corrected by two adaptive corrections. Finally, MannekenPix provides a statistical estimate of the uncertainty affecting the automatic picks following a calibration derived from reference picks and weights provided by the user.

Chapter 4 presents the case study of the application of MannekenPix to the local data set gathered for the Dead Sea region. In order to allow potential users of MannekenPix to better evaluate the profile of the data, a detailed characterization of the data set is presented. The picking hit rate and accuracy are examined next, followed by the estimated picking uncertainties. Residual times from a simple relocation conclude Chapter 4.

The conclusions of the work are presented in Chapter 5, along with the main results from the application of MannekenPix to the seismicity of Italy.

Chapter 2

Lower-crustal strength under the Dead Sea basin from local earthquake data and rheological modeling

F. Aldersons, Z. Ben-Avraham, A. Hofstetter, E. Kissling and T. Al-Yazjeen

Earth and Planetary Science Letters (2003) 214 129-142

ABSTRACT

We studied the local seismicity of the Dead Sea basin for the period 1984-1997. Sixty percent of well-constrained microearthquakes ($M_L \leq 3.2$) nucleated at depths of 20-32 km and more than 40 percent occurred below the depth of peak seismicity situated at 20 km. With the Moho at 32 km, the upper mantle appeared to be aseismic during the 14-year data period. A relocation procedure involving the simultaneous use of three regional velocity models reveals that the distribution of focal depths in the Dead Sea basin is stable. Lower-crustal seismicity is not an artifact created by strong lateral velocity variations or data-related problems. An upper bound depth uncertainty of ± 5 km is estimated below 20 km, but for most earthquakes depth mislocations should not exceed ± 2 km. A lithospheric strength profile has been calculated. Based on a surface heat flow of 40 mWm^{-2} and a quartz-depleted lower crust, a narrow brittle to ductile transition might occur in the crust around 380°C at a depth of 31 km. For the upper mantle, the brittle to ductile transition occurs in the model at 490°C and at 44 km depth. The absence of micro-seismicity in the upper mantle remains difficult to explain.

2.1 INTRODUCTION

It is commonly assumed (Meissner and Strehlau, 1982; Chen and Molnar, 1983) that no significant seismicity exists in rifts at depths greater than 15-20 km. The lower crust in continental extension zones is indeed often aseismic due to elevated temperatures accompanying lithospheric thinning. In relation to a more recent trend of research (Cloetingh and Burov, 1996; Maggi et al., 2000; Jackson, 2002), microearthquakes from Israel and Jordan suggest that the

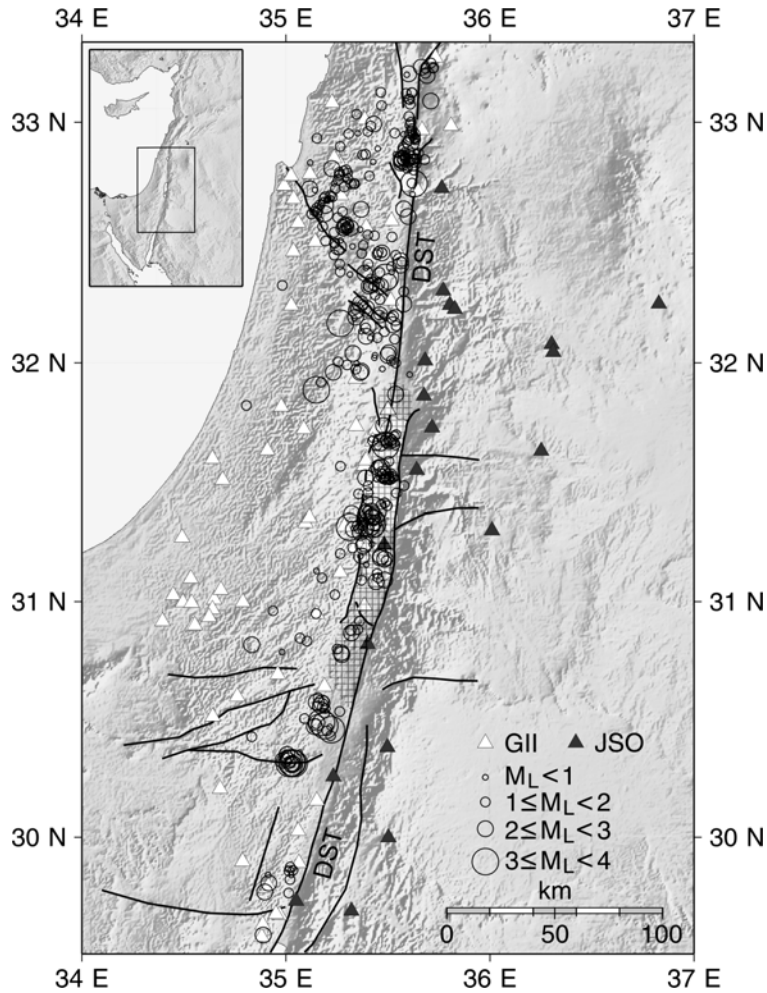


Figure 2.1 The Dead Sea region. Topographic shaded relief derived from Globe 1-km elevation data. The square grid fill defines the Dead Sea basin. Seismicity: 410 well-constrained earthquakes (1984-1997) recorded by short-period stations (triangles) of GII (Israel) and JSO (Jordan). DST: Dead Sea Transform.

lower crust (20-32 km) has probably kept a significant strength under the Dead Sea basin. A similar behavior of the lower crust in continental extension regimes has already been observed, mainly in the western branch of the East African rift system (Shudofsky et al., 1987; Nyblade and Langston, 1995; Camelbeek and Iranga, 1996) and in the Baikal rift system (Déverchère et al., 2001). Anomalous deep crustal earthquakes are also known to occur in the western United States (Wong and Chapman, 1990; Bryant and Jones, 1992) and below the northern Alpine foreland of Switzerland (Deichmann and Rybach, 1989). Focal depths of earthquakes are generally difficult to constrain tightly. The most reliable estimates are usually those derived from local earthquake data if well-distributed stations are present and a good knowledge of the velocity field has been gained. The Dead Sea basin and its surroundings (Figure 2.1) have been monitored locally for more than 10 years by permanent digital seismic stations. In addition, the

region has been the focus of detailed geophysical studies during past decades. These studies include bathymetry (Neev and Hall, 1979), seismic reflection (Neev and Hall, 1979; ten Brink and Ben-Avraham, 1989), seismic refraction (Ginzburg et al., 1981; Ginzburg and Ben-Avraham, 1997), gravity (ten Brink et al., 1990; ten Brink et al., 1993), magnetism (Frieslander and Ben-Avraham, 1989; Al-Zoubi and Ben-Avraham, 2002), heat flow (Ben-Avraham et al., 1978) and seismicity (van Eck and Hofstetter, 1989, 1990).

Constraints provided by these studies not only lead to more robust results, they also allow the development of geophysical models better tuned to the geological environment. In this article we present evidence from local earthquake data and rheological modeling that lead us to acknowledge the existence of a significant lower-crustal strength under the Dead Sea basin.

2.2 GEOLOGICAL SETTING

The Dead Sea Transform (Figure 2.1) is an intracontinental plate boundary resulting from the late-Cenozoic breakup of the Arabo-African continent. This boundary extends over 1,000 km from the zone of sea floor spreading at the southern tip of the Sinai Peninsula to the Taurus-Zagros zone of convergence in Turkey (Freund, 1965). The Dead Sea basin is an active pull-apart (Quennell, 1958) located along the Dead Sea Transform. The amount of left-lateral motion along the Transform in the Dead Sea region is estimated at 105 km (Quennell, 1958; Freund et al., 1970). Motion along the Transform started around 20 Ma ago during Miocene times (Bartov et al., 1980). Although the Dead Sea basin originated at the beginning of the Transform motion, it did not develop into a topographic low before early Pliocene (Garfunkel, 1997). Pliocene sediments are mostly composed of evaporites, among which halite is the main constituent. Deformation of Pliocene salt has created numerous diapirs in the basin (Neev and Hall, 1979). During Pleistocene times, basin subsidence accelerated (ten Brink and Ben-Avraham, 1989) and allowed the accumulation of several kilometers of lacustrine clastics, carbonates and evaporites. Today the basin is a morphotectonic depression over 130 km long and 7-18 km wide (Figure 2.1). It is a seismically active section (van Eck and Hofstetter, 1989, 1990) along the Dead Sea Transform, for which some 4,000 years of combined archaeological, historical and instrumental seismicities are documented (Ben-Menahem, 1991). For the northern half of the Dead Sea basin (main lake and salt pans), earthquakes of $M_L \geq 5.8 \pm 0.2$ have a recurrence interval of approximately 160 years (Shapira, 1997). The last earthquake of such a magnitude (estimated magnitude 6.2 by Turcotte and Arieh, 1988) occurred in the main lake (Shapira et al., 1993) in 1927 at an unknown depth.

2.3 SEISMICITY

Figure 2.2 displays a depth section of seismicity along the Dead Sea Transform between Aqaba-Elat in the south and the Sea of Galilee in the north for the period 1984-1997. Out of 2,283 routinely recorded events, a first selection resulted in 653 events with at least 8 P readings each and an azimuthal gap smaller than 180 degrees. To further exclude events with poorly constrained depths, an empirical criterion based on the epicentral distance to the nearest station (D_{min}) was adopted for events with an estimated depth of less than 21 km (Figure 2.3). For depths greater than 21 km, all events from the first selection appeared as valid candidates and were included. Two hundred forty-three events from the first selection of 653 events were rejected while the remaining were included and are the 410 earthquakes displayed in Figure 2.1 and Figure 2.2. Our study focuses on the Dead Sea basin and discusses whether the apparently unusual focal depths observed there are part of a broader phenomenon or represent a local anomaly.

We started the Dead Sea seismicity study by a series of relocations aimed at detecting flaws like unstable local minima solutions. Most first P arrivals from earthquakes nucleating in the basin were carefully repicked manually, and weighted according to the quality of their onsets. Forty-two well-constrained earthquakes qualified (Figure 2.4a) and relocations were performed with program Velest (Kissling, 1994).

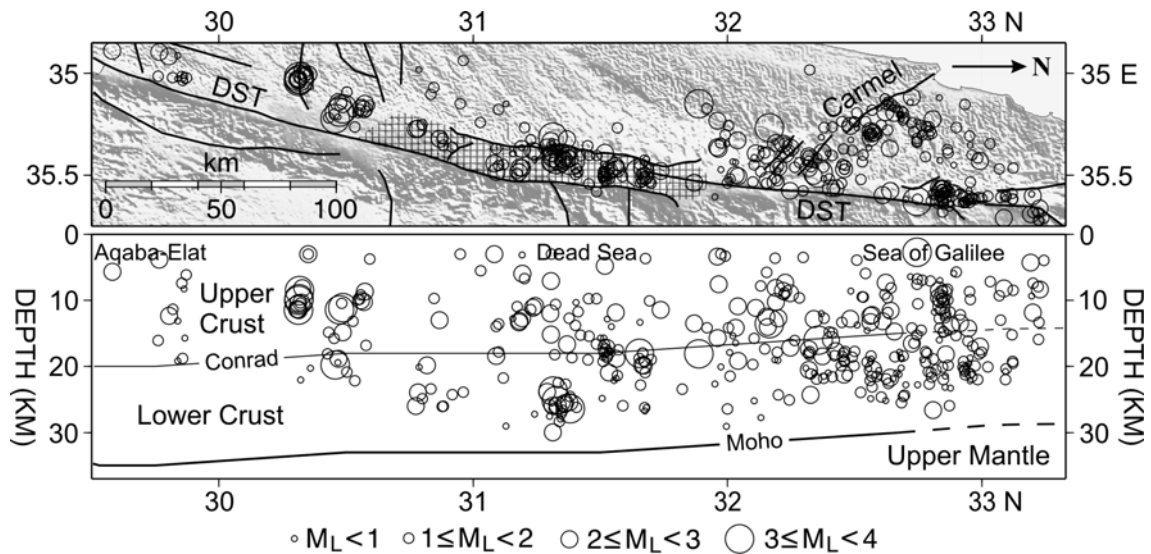


Figure 2.2 Depth section of well-constrained seismicity (410 earthquakes, 1984-1997) along the Dead Sea Transform (DST) from Aqaba-Elat to the Sea of Galilee. The square grid fill defines the Dead Sea basin on the map. Conrad and Moho discontinuities from Ginzburg et al. (1981).

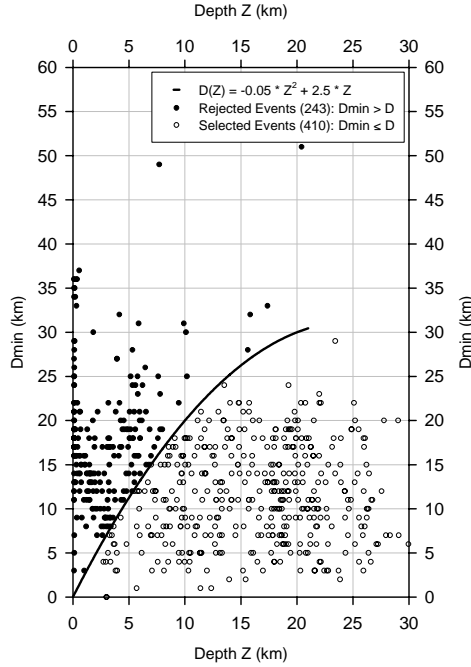


Figure 2.3 Near-epicentral distance selection for 653 events from Aqaba-Elat to the Sea of Galilee ($P \geq 8$ and Gap $< 180^\circ$). Z is the estimated hypocentral depth and D_{min} is the epicentral distance to the nearest station.

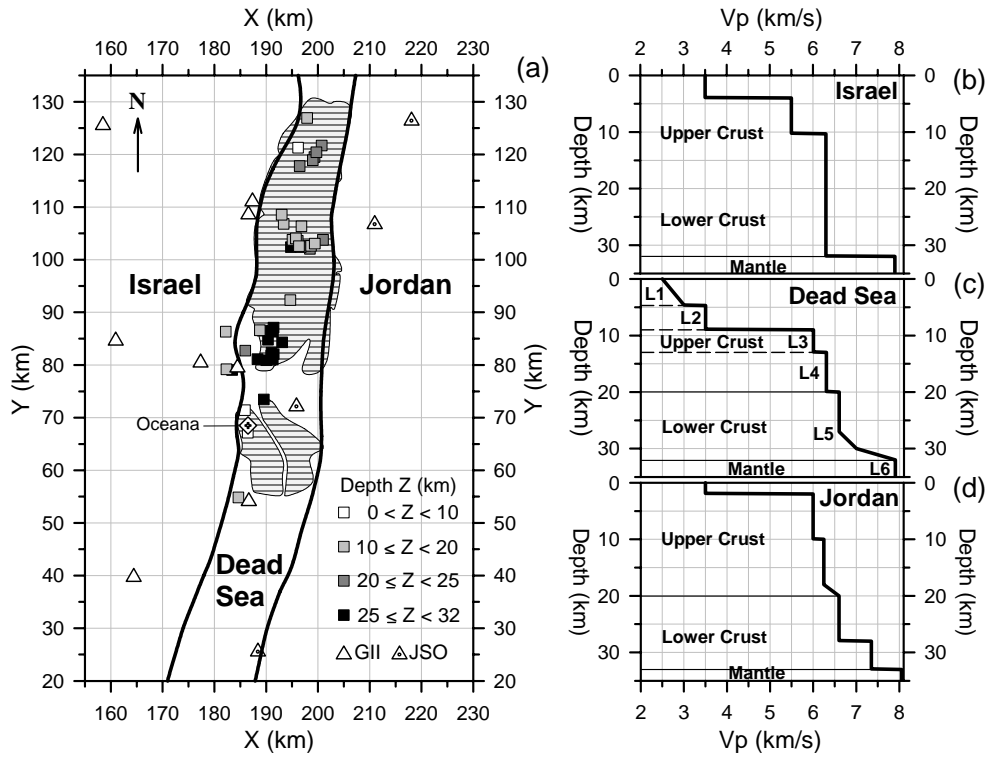


Figure 2.4 (a) Structural units and stations (triangles) in the Simulps relocation. Epicenters and depths of the set of 42 earthquakes (squares) relocated with Velest. (X, Y) are Rectangular Israel Grid coordinates (Survey of Israel). Oceana is a salt quarry. (b,c,d) The three velocity models used in the Simulps relocation.

Model Israel (Figure 2.4b) was applied to the whole volume of earthquakes and stations. This model is similar to the routine model used by the Geophysical Institute of Israel (GII) but we derived it from quarry blasts located in the vicinity of the Dead Sea basin. Several relocations were conducted with Velest during which the effects of various initial conditions were explored. During these relocations, it was not considered a significant problem if the focal depths of a few earthquakes shifted up or down. Our main criterion was the stability of the distribution itself, not the perfect stability of every estimate. Tests like these are usually good at detecting flaws in hypocentral parameters like local minima solutions. No such flaws were ever observed. Depths derived from model Israel are stable. The distribution of depths in the Dead Sea basin is plotted in Figure 2.5a. It shows that 60 percent of well-constrained microearthquakes ($M_L \leq 3.2$) nucleated at depths of 20-32 km and that more than 40 percent occurred below the depth of peak seismicity situated at 20 km. The upper mantle appeared to be aseismic during the 14-year data period.

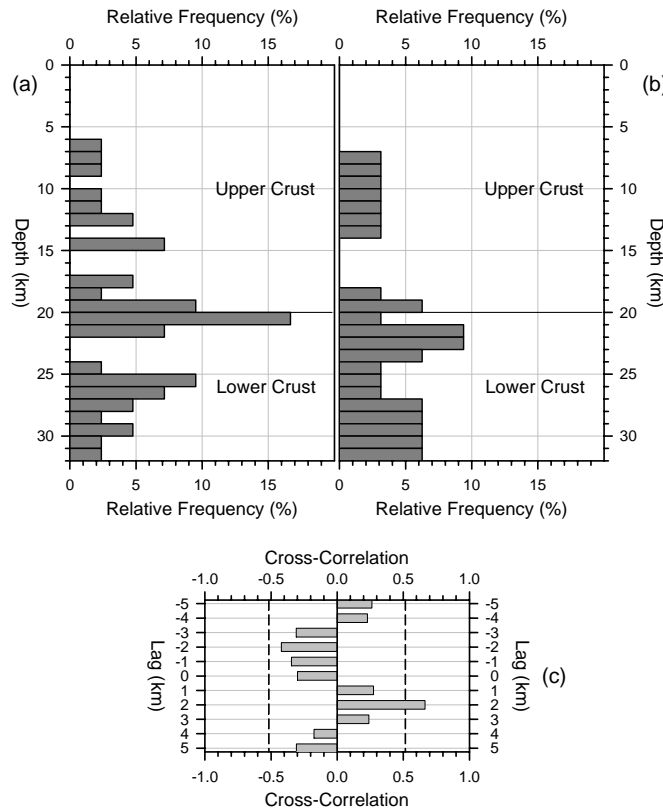


Figure 2.5 Focal depths of well-constrained earthquakes in the Dead Sea basin. (a) Depth distribution of 42 earthquakes located with regional velocity model Israel (Figure 2.4b) in Velest. (b) Depth distribution of 32 earthquakes located with three regional velocity models in Simulps (Figure 2.4). (c) Cross-correlation of (a) and (b) between 17 and 32 km depth. The dashed lines delineate the 95 % confidence interval assuming (a) and (b) to be uncorrelated. A significant correlation of 0.66 is found at Lag = + 2 km.

Although depths derived from model Israel are stable, velocities in the uppermost 9 km of the model are too high for the Dead Sea basin (compare model Israel versus model Dead Sea, Figure 2.4b and Figure 2.4c). This might create an artificial deepening of focal depths for earthquakes nucleating in the basin. In order to evaluate this possibility, program Simulps (Thurber, 1984) was also used to relocate earthquakes.

In each of the three structural units of Figure 2.4a, a one-dimensional velocity model was defined. The boundaries between the three units were derived from the maximum gradient of gravity (ten Brink et al., 1993). All three models are rather similar below the depth of 10 km. Models Israel and Jordan (El-Isa, 1990; JSO, 1993) are rather similar throughout. The first layer is however thinner in model Jordan compared to model Israel and no contrast between the upper and lower crust is present in model Israel. The upper crust in the Dead Sea velocity model was derived from a refraction profile in the north basin of the Dead Sea (Ginzburg and Ben-Avraham, 1997), while the lower crust was derived from a deep refraction experiment by Ginzburg et al. (1981). From the abrupt termination of the cusp (Bullen, 1960; Ginzburg et al., 1979), the authors infer the existence of a 5-km transition zone between the main lower-crustal velocity of 6.6 km/s and the upper mantle velocity of 7.9 km/s, a feature apparently missing in adjacent areas outside the rift. According to the same study, the Dead Sea basin is underlain by a Moho at about 32 km depth, a value supported by gravity modeling (ten Brink et al., 1993).

Earthquakes were relocated by Simulps according to the velocities and boundaries of the three structural units. No attempt was made to derive optimal velocities from the data themselves. The selection of a fine discretization grid resulted in restrictions on the maximum size of the model. Due to the smaller number of stations available, only 6 P arrivals were required for qualifying events. Among the 42 earthquakes relocated by Velest, 32 qualified for the Simulps relocation. Epicentral distances to stations never exceeded 85 kilometers and no more than 11 arrivals per earthquake could be observed. Most seismic rays were confined to the upper crust and rays from deeper earthquakes quickly turned up. Figure 2.5b displays the depth distribution using Simulps. The lower-crustal seismicity remains in place and the whole distribution displays features similar to those in the distribution computed with Velest, although deeper by 2 km in the lower half. Nearly identical results were obtained without the transition zone at the base of the lower crust. One should not however conclude that the Simulps relocation is more reliable than the Velest relocation. The purpose of the Simulps relocation was to test the possible influence of slow velocities in the upper section of the Dead Sea basin. The results show that the lower-crustal seismicity detected by Velest is not an artifact created by these slow basinal velocities.

Two independent evaluations of the uncertainty on depths derived from model Israel have been made. First, we applied perturbations to model Israel and relocated the 42 well-constrained earthquakes with resulting models in Velest. This method evaluates the sensitivity of depths to departures from model Israel, and it provides individual error bars for each of the 42 earthquakes. Velocity perturbations of $\pm 5\%$, $\pm 10\%$, and depth perturbations of $\pm 10\%$, $\pm 20\%$, were added to model Israel (Table 2.1 and Figure 2.6a). Velocity perturbations of $+10\%$ for layer 2 and -10% for layer 3 were not applied because they would have created some unrealistic velocity inversions with depth from layer 2 to layer 3. Consequently and in order to preserve symmetry around model Israel, velocity perturbations of -10% for layer 2 and $+10\%$ for layer 3 were also not applied. The bottom depth of layer 3 (the Moho) was not perturbed by more than $\pm 9\%$ following gravity results implying that the Moho should not be anomalously elevated by more than 2-3 km under the Dead Sea basin (ten Brink et al., 1990; ten Brink et al., 1993). Perturbations of ± 0.250 km/s around the upper mantle velocity of 7.9 km/s appeared also to be quite sufficient. All combinations of perturbations defined in Table 2.1 represent a set of 16,875 velocity models.

Depths and error bars derived from this set are plotted in Figure 2.6b. Two main conclusions emerge. First, according to the 5th to 95th percentile intervals, at least 25 % of well-constrained microearthquakes under the Dead Sea basin can belong neither to the upper-crustal seismicity nor to a hypothetical upper-mantle seismicity. Second, error bars tend to narrow for deeper earthquakes (Figure 2.7), expressing a lesser sensitivity to model perturbations in the lower crust compared to the upper crust.

Table 2.1 Perturbations to model Israel

Layer	V_p (km/s)	V_p perturbation (%)	Layer bottom (km)	Bottom perturbation (%)
Layer 1	3.150	-10.0 %	3.280	-20.0 %
	3.325	-5.0 %	3.690	-10.0 %
	3.500	0.0 %	4.100	0.0 %
	3.675	+5.0 %	4.510	+10.0 %
	3.850	+10.0 %	4.920	+20.0 %
Layer 2			8.640	-20.0 %
	5.225	-5.0 %	9.720	-10.0 %
	5.500	0.0 %	10.800	0.0 %
	5.775	+5.0 %	11.880	+10.0 %
			12.960	+20.0 %
Layer 3			29.000	-9.1 %
	5.985	-5.0 %	30.450	-4.5 %
	6.300	0.0 %	31.900	0.0 %
	6.615	+5.0 %	33.350	+4.5 %
			34.800	+9.1 %
Layer 4	7.650	-3.2 %	Half-space	
	7.900	0.0 %		
	8.150	+3.2 %		

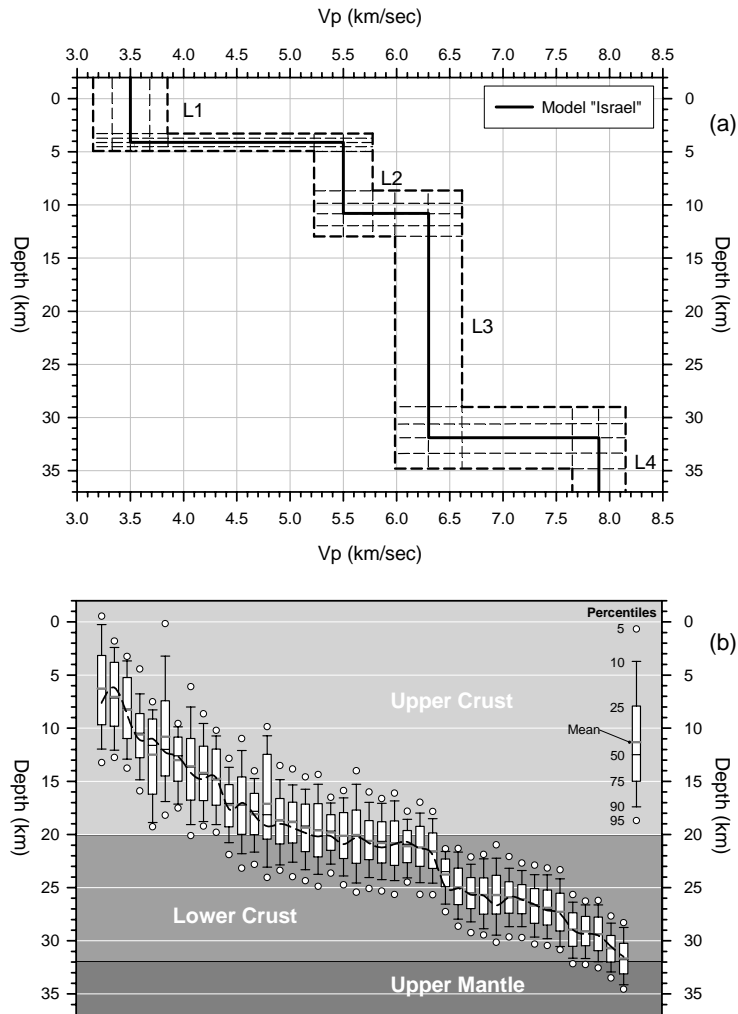


Figure 2.6 (a) Depth and velocity perturbations defining 16,875 models around model Israel. L1 to L4 are the four layers of model Israel. (b) Hypocentral depths and error bars for the 42 earthquakes relocated by Velest with perturbed models defined in (a). Earthquakes sorted according to increasing median (50th percentile) depth. The black dashed line marks depths obtained with model Israel.

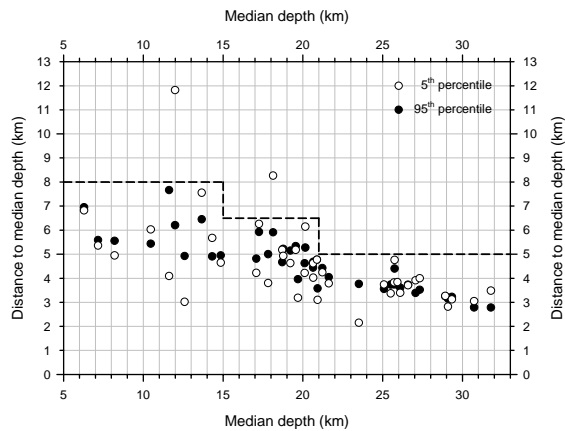


Figure 2.7 Upper-bound to the uncertainty on depths (dashed line) from perturbations to model Israel.

This is a coherent behavior of a distribution of depths since seismic rays from shallower earthquakes tend to propagate twice (by refraction) through more layers than deeper earthquakes do. Error bars show also that several lower-crustal earthquakes close to 20 km depth could have nucleated in the upper crust but the reverse is also true, a similar number of upper-crustal earthquakes could have nucleated in the lower crust. The net result is then very close to the original distribution of depths. Close to the Moho, only the two deepest earthquakes are realistic candidates for an upper-mantle origin but model Israel locates them in the lower crust. From Figure 2.7, an upper-bound to the uncertainty in the range of depths 21-32 km is then ± 5 km.

As a second approach to depth uncertainties, we determined true depth errors for a series of blasts from quarry Oceana. Oceana is a Dead Sea salt quarry located on the western salt pan south of the main lake and only 15 km away from the main cluster of deepest earthquakes (Figure 2.4a). We only considered well-constrained blasts with at least 8 P readings, and an azimuthal gap not greater than 150 degrees. In addition, we required Dmin (distance of closest station to the estimated epicenter) to be 3-4 km and we rejected events not explicitly attributed to Oceana by the GII analysts. Figure 2.8 shows that the number of P readings for the 24 Oceana blasts is lower than for the 42 earthquakes while gaps are similar. It shows also that earthquakes below 20 km depth have all a Dmin value lower than the estimated depth, an important feature regarding depth control. As a whole, Oceana blast data are then not as well-constrained as our set of earthquakes.

Nevertheless, Figure 2.9b reveals that Oceana blasts located with model Israel do not display true depth errors greater than 2 km, with the exception of two outliers. Since blast locations are generally less accurate than earthquake locations in a given area due to poorly approximated refracted rays near sources at the surface (Kissling, 1988), even smaller depth mislocations are possible for earthquakes in the basin.

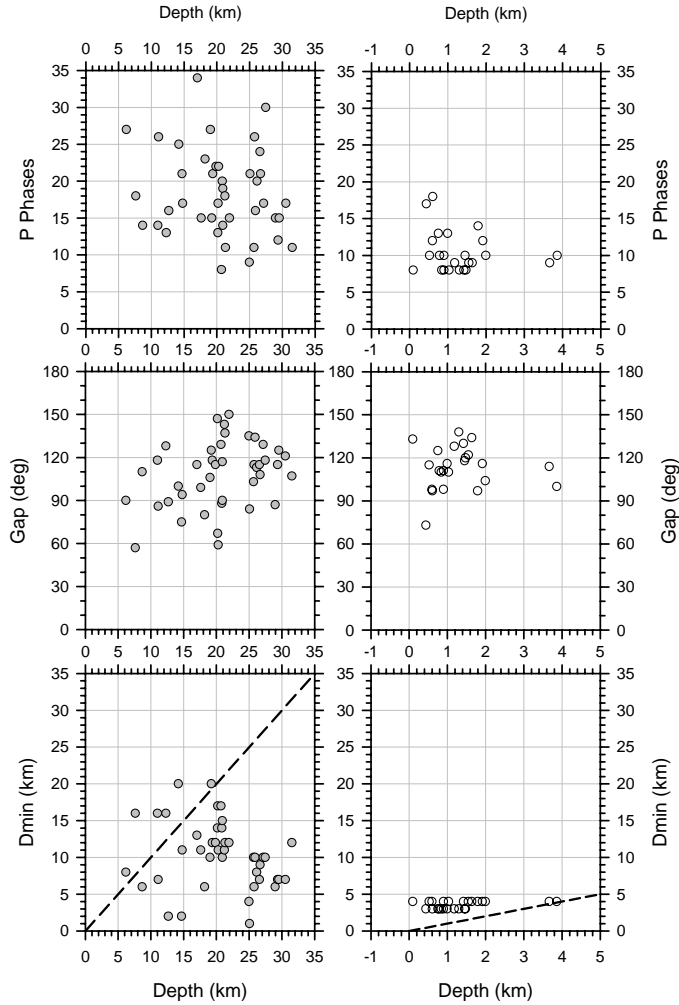
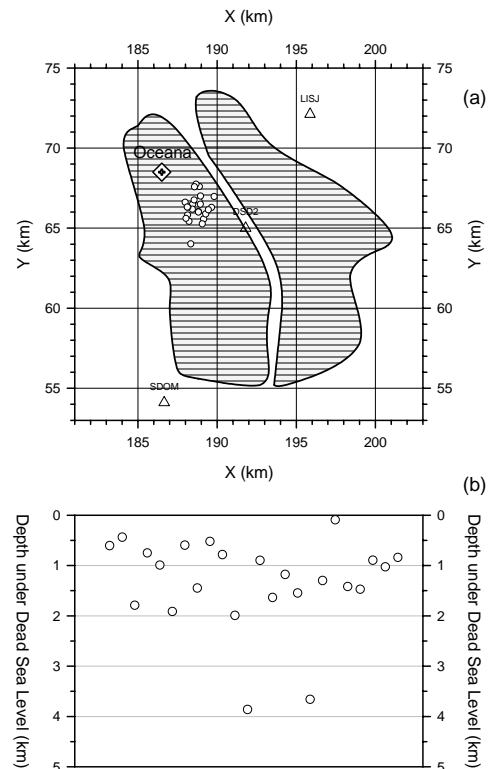


Figure 2.8 (Left) Number of P phases, azimuthal gap and Dmin for the 42 earthquakes located with model Israel in Velest. (Right) Same as (Left) but for 24 blasts of quarry Oceana. Dmin is the distance of the closest station to the estimated epicenter. The dashed line outlines values for which Dmin is equal to the estimated depth.

Figure 2.9 (a) Epicentral locations (model Israel in Velest) of 24 Oceana quarry blasts. X and Y are Cassini-Soldner (Survey of Israel) projection coordinates. Epicenters as white circles. (b) Depth error of blasts plotted in (a).



2.4 RHEOLOGY

A lithospheric strength profile (e.g. Ranalli, 1987) has been calculated using the five crustal layers defined in velocity model Dead Sea (Figure 2.4c) and one additional layer for the upper mantle. Below 20 km, the lower crust was modeled as diabase, a lithology consistent with the main value of 6.6 km/s (Christensen and Mooney, 1995) in layer 5 of the velocity model. The transition zone at the base of the crust was discarded.

We derived the geotherm for the Dead Sea basin from the one-dimensional equilibrium heat conduction (Turcotte and Schubert, 2002) equation

$$\frac{\partial^2 T}{\partial z^2} = -\frac{A}{k} \quad (2.1)$$

where T is the temperature, z the depth, A the radiogenic heat generation rate per unit volume, and k the thermal conductivity. This second-order differential equation can be solved given two boundary conditions. Assuming the surface to be at $z = 0$ and z increasing with depth, the surface heat flow and surface temperature as boundary conditions are respectively $Q = -k\partial T / \partial z = -Q_0$ at $z = 0$, and $T = T_0$ at $z = 0$. The surface heat flow $Q = -Q_0$ is negative because heat is flowing upwards and z is positive downwards. For a layered model where each layer i has a given thermal conductivity k_i and a radiogenic heat generation rate A_i , we obtain the temperature $T_{n,z}$ in layer n at a depth z from the surface:

$$T_{n,z} = -\frac{A_n}{2k_n}z^2 + \left[\sum_{i=1}^{n-1} \left(\frac{A_{i+1}}{k_{i+1}} - \frac{A_i}{k_i} \right) z_i + \frac{Q_0}{k_1} \right] z + \sum_{i=1}^{n-1} \left(\frac{A_i}{2k_i} - \frac{A_{i+1}}{2k_{i+1}} \right) z_i^2 + T_0 \quad (2.2)$$

where z_i is the bottom depth of layer i . Using six layers and average parameters, the one-dimensional geotherm is only sensitive to two variables: the surface heat flow, and the thermal conductivity in the first layer. These two variables exert a great influence on the slope of the geotherm at the surface and therefore play a major role on the estimated temperature in the deeper part of the model. The average measured heat flow in the northern Dead Sea basin is 38 mWm⁻² (Ben-Avraham et al., 1978; Ben-Avraham, 1997) and it is 42 mWm⁻² (Eckstein and Simmons, 1978) west of the basin. These values are very similar to the uniform heat flow measured in the eastern Mediterranean (Erickson, 1970). Consequently, the surface heat flow Q_0 of 40 ± 2 mWm⁻² we used to compute the equilibrium geotherm appears to be well constrained. The surface temperature T_0 was set in the model at 5 °C. Regarding the first layer, Plio-Pleistocene to recent sediments are composed of fluvial and lacustrine clastics, marls, chalks and evaporites. The few quaternary conductivities available in the Dead Sea basin (Eckstein and Simmons, 1978) are very low (1.25 Wm⁻¹K⁻¹) but were measured through only the first 150 m

from the surface. In order to get a brittle to ductile transition around 30 km with a surface heat flow of 40 mWm^{-2} , a mean thermal conductivity of at least $2.1 \text{ Wm}^{-1}\text{K}^{-1}$ is required in the first layer. This figure appears as a realistic value taking into account compaction and a content of 5-10 percent of highly conductive evaporites in the first layer. It is unlikely that the true value over the thickness (4.7 km) of the first layer would be lower than $2.0 \text{ Wm}^{-1}\text{K}^{-1}$.

The radiogenic heat generation rate is poorly constrained in the Dead Sea region. However, about 110 m of recent sediments from two cores (Stiller et al., 1985) taken in the Dead Sea lake contained on the average 3.87 ppm of ^{238}U , 4.15 ppm of ^{232}Th and 1.1 ppm of ^{40}K . These U and Th values are within the normal range of sedimentary materials found worldwide (Rodgers and Adams, 1969). For layers 1 and 2 of density $2,150 \text{ kgm}^{-3}$ and $2,550 \text{ kgm}^{-3}$, the total heat generation rate A is then (Beardsmore and Cull, 2001) 1.0 and $1.2 \text{ }\mu\text{Wm}^{-3}$ respectively. For the crustal plutonic rocks of the model, we used the relations between the velocity of P waves V_p and the heat generation rate A of Rybach and Buntebarth (1984). With their relation for Phanerozoic rocks, we obtained (at 200 MPa) 2.27 and $1.27 \text{ }\mu\text{Wm}^{-3}$ for layer 3 (granite) and layer 4 (quartz diorite) respectively. To account for the fact that the last stage of major regional plutonism occurred (Garfunkel, 1999) during late Proterozoic (Pan-African orogeny), we applied an arbitrary correction of -20 % to the computed values. The adopted heat generation rates for layers 3 and 4 are then 1.8 and $1.0 \text{ }\mu\text{Wm}^{-3}$ respectively. For layer 5 (diabase), the Rybach-Buntebarth relation for Precambrian rocks led to a value of $0.2 \text{ }\mu\text{Wm}^{-3}$. An average value of $0.007 \text{ }\mu\text{Wm}^{-3}$ (Ranalli, 1987) was adopted for the upper mantle.

A frictional (brittle) failure criterion following Byerlee's law of friction (Byerlee, 1968, 1978) can be expressed from Anderson's theory of faulting (Anderson, 1951) as (Sibson, 1974; Ranalli, 1987)

$$\sigma_1 - \sigma_3 = \alpha \rho g z (1 - \lambda) \quad (2.3)$$

where $\sigma_1 - \sigma_3$ is the failure stress, α a parameter depending on the type of faulting, ρ the average density, g the acceleration of gravity, z the depth, and λ is the ratio of pore fluid pressure to lithostatic pressure. For a friction coefficient of 0.75, values of the fault parameter α are 3.0, 1.2 and 0.75 for thrust, strike-slip and normal faults respectively (Sibson, 1974; Ranalli, 1987). We adopted an intermediate value α of 1.05 between strike-slip and normal faulting as earthquake focal mechanisms (van Eck and Hofstetter, 1989) and seismic reflection profiles (ten Brink and Ben-Avraham, 1989) show that both types of faults coexist in the basin. The average density ρ in the model is $2,670 \text{ kgm}^{-3}$ for the crust, and $3,370 \text{ kgm}^{-3}$ (Turcotte and Schubert, 2002) for the upper mantle above the 220 km discontinuity. The pore fluid pressure was set at an intermediate

value between hydrostatic and dry. Equation 2.3 provides a lower limit to the failure stress because it assumes an ideal orientation of faulting with respect to the stress field.

Ductile deformation in the crust is dominated by power-law creep (dislocation climb) given (e.g. Ranalli, 1987) by

$$\sigma_1 - \sigma_3 = \left(\frac{\dot{\epsilon}}{D} \right)^{\frac{1}{n}} e^{\frac{E_C}{nRT}} \quad (2.4)$$

where $\sigma_1 - \sigma_3$ is the flow stress for a given strain rate, $\dot{\epsilon}$ the strain rate, D the Dorn parameter, n the stress exponent, E_C the creep activation energy, R the gas constant, and T is the temperature. Values for D , n and E_C are provided in Table 2.2.

Table 2.2 Dislocation climb (power-law) creep parameters

Rock / Mineral	Layer	D (Pa ⁻ⁿ s ⁻¹)	n	E_C (kJ mol ⁻¹)
Quartzite (wet) ¹	1, 2	3.99×10^{-18}	2.3	154
Granite ¹	3	7.92×10^{-29}	3.2	123
Quartz diorite ²	4	5.02×10^{-18}	2.4	219
Diabase ³	5	8.04×10^{-25}	3.4	260
Olivine ⁴	6	7.00×10^{-14}	3.0	511

D is the Dorn parameter, n the stress exponent and E_C is the creep activation energy.

¹ From Ranalli and Murphy (1987), data from compilation of Kirby (1983).

² From Kirby (1983), data from Hansen and Carter (1982).

³ From Kirby (1983), data from Shelton and Tullis (1981).

⁴ From Goetze (1978).

In the mantle, ductile deformation is controlled by the power-law creep (equation 2.4) of olivine only for deviatoric stresses below 200 MPa. For deviatoric stresses greater than 200 MPa, the glide-controlled creep of olivine is assumed to dominate (Goetze, 1978; Kameyama et al., 1999):

$$\sigma_1 - \sigma_3 = \sigma_0 \left(1 - \left[\frac{RT}{E_p} \ln \left(\frac{\dot{\epsilon}_0}{\dot{\epsilon}} \right) \right]^{\frac{1}{2}} \right) \quad (2.5)$$

In equation 2.5, σ_0 is the Peierls reference stress, E_p the activation energy for the Peierls mechanism, and $\dot{\epsilon}_0$ is the reference strain rate. Values for σ_0 , E_p and $\dot{\epsilon}_0$ are provided in Table 2.3.

Table 2.3 Dislocation glide creep parameters

Rock / Mineral	Layer	σ_0 (Pa)	$\dot{\epsilon}_0$ (s ⁻¹)	E_p (kJ mol ⁻¹)
Olivine	6	8.5×10^9	5.7×10^{11}	535

σ_0 is the Peierls reference stress, E_p the activation energy for the Peierls mechanism, and $\dot{\epsilon}_0$ is the reference strain rate. Data from Goetze (1978).

A strain rate $\dot{\epsilon}$ of $2 \times 10^{-15} \text{s}^{-1}$ was used in the model. This value was derived from a relative plate motion of 5 mm/yr, an intermediate value between short-term global positioning results of 2.6 ± 1.1 mm/yr (Pe'eri et al., 2002) and long-term geological estimates of 6-10 mm/yr (Quennell, 1958; Freund, 1965).

At a given depth, the smaller value of the failure stress and the flow stress gives the yield strength. A rheological profile (strength envelope) is a curve of yield strength versus depth. At low temperatures, the failure stress is lower than the flow stress and brittle behavior dominates the deformation of rocks. The opposite is true at high temperatures where aseismic ductile behavior prevails. The usually sharp brittle to ductile transitions in the lithosphere are supposed to correspond to boundaries of seismogenic zones. Several brittle to ductile transitions and related seismogenic zones can occur in the lithosphere. The interpretation of strength profiles is based on rather strong assumptions (Scholz, 2002) but the method provides a useful first-order estimate of the rheology.

Figure 2.10 displays the results with a surface heat flow of 40 mWm^{-2} and a thermal conductivity of $2.1 \text{ Wm}^{-1}\text{K}^{-1}$ in the first layer. A narrow brittle to ductile transition occurs in the crust around 380°C at 31 km depth. With a heat flow of 42 mWm^{-2} and a conductivity of $2.0 \text{ Wm}^{-1}\text{K}^{-1}$ in the first layer, the transition would occur around 390°C at a depth of 26 km. However, parameters we adopted from Shelton and Tullis (1981) for diabase in the lower crust do not represent dry deformation conditions, as their samples were only dried at 160°C . A recent experimental study on the creep of diabase (Mackwell et al., 1998) revealed that this type of rock is significantly stronger under truly dry conditions. Consequently, if the lower crust under the Dead Sea basin has a basaltic composition and is dry, it should even be ductilely stronger than what we report. In that case, no brittle to ductile transition would occur at all in the lower crust except for unrealistic Moho temperatures above 600°C and related surface heat flows greater than 55 mWm^{-2} . In the upper mantle, the brittle to ductile transition occurs in the model at 44 km depth and at 490°C but as said earlier, the upper mantle appears to be aseismic during the 14-year data period.

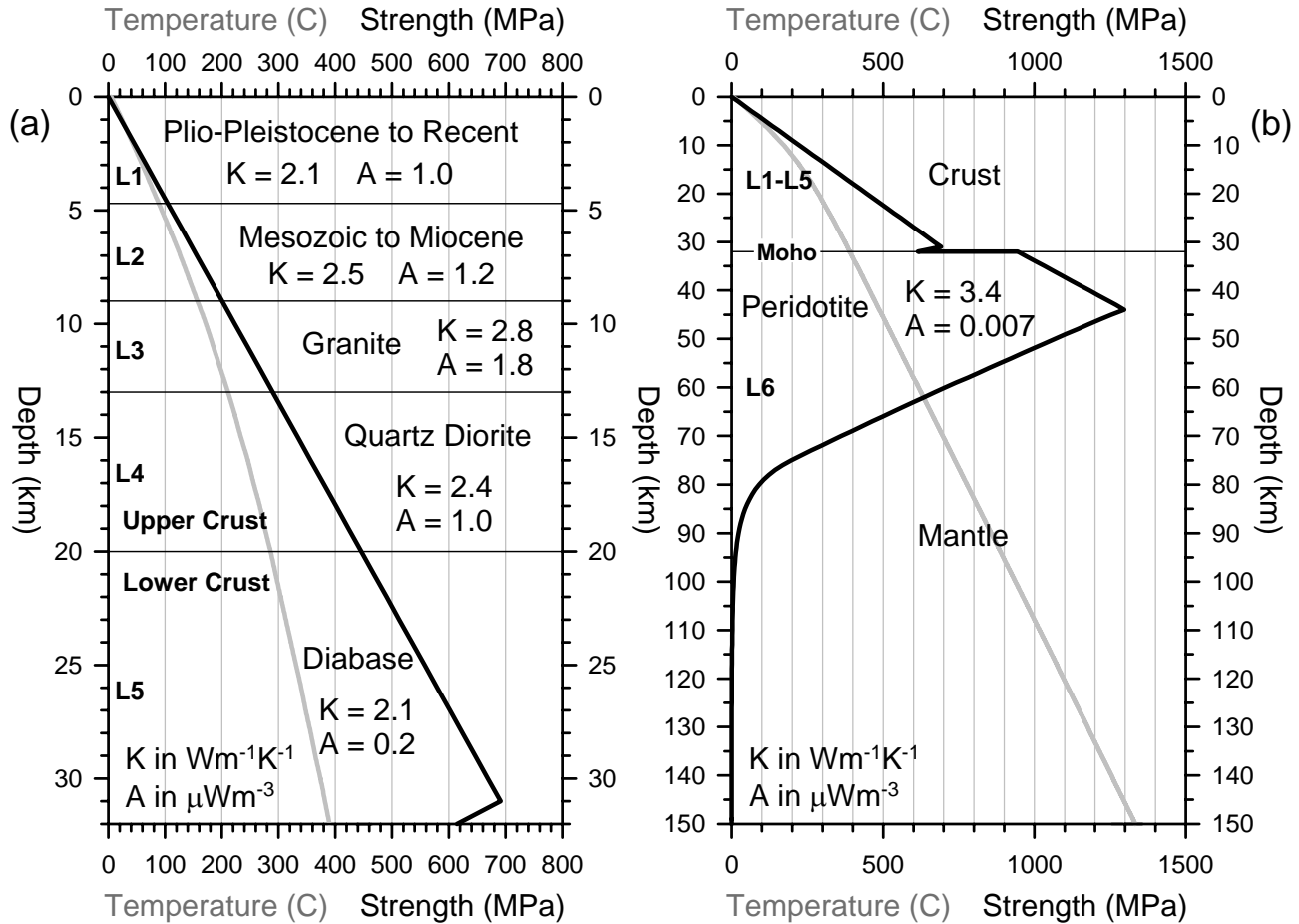


Figure 2.10 Rheology of the Dead Sea basin. (a) Crustal geotherm (gray) and strength profile (black). (b) Lithospheric geotherm (gray) and strength profile (black). Surface heat flow of 40 mWm^{-2} . K is the thermal conductivity and A is the radiogenic heat production rate.

2.5 DISCUSSION AND CONCLUSIONS

In the Dead Sea basin, well-constrained microearthquakes ($M_L \leq 3.2$) display continuous focal depths down to the Moho located at a depth of 32 km. Sixty percent of these well-constrained microearthquakes nucleated at depths of 20-32 km and more than 40 percent occurred below the depth of peak seismicity situated at 20 km. An upper bound uncertainty of ± 5 km is estimated for depths greater than 20 km, but depth mislocations should not exceed ± 2 km for most earthquakes. No data support the view that lower-crustal seismicity could be nothing more than an artifact due to earthquakes actually nucleating in the upper crust or in the upper mantle. Due to uncertainties on focal depths, a few earthquakes might well have nucleated in the upper mantle but a distinct population of upper mantle earthquakes is missing.

Most parameters in rheological modeling are commonly poorly constrained, and the method itself suffers from inherent limitations (Scholz, 2002). Therefore, strength envelopes generally cannot serve to accurately estimate the thickness of the seismogenic zone. In the case of the Dead Sea basin, we are merely able to say that a brittle to ductile transition around 31 km depth (where semi-brittle failure would prevail) is consistent with a surface heat flow of 40 mWm^{-2} . The low value of the regional surface heat flow is a good indication that the lower crust might be cool and brittle. Nevertheless, since conductive thermal anomalies can take millions of years to reach the surface, the surface heat flow does not always reflect lower-crustal temperatures. With such a clear lower-crustal seismicity as the one monitored in the Dead Sea basin, a significant departure from the equilibrium geotherm is however not expected. It appears therefore valid to use the equilibrium geotherm to extrapolate temperatures to lower-crustal depths. The estimated temperature of 380°C at the brittle to ductile transition is clearly below the temperature of 450°C ($350 \pm 100^\circ\text{C}$) generally considered (Chen and Molnar, 1983) as the limiting temperature of the deepest crustal seismogenic zone. In addition, a crustal seismic investigation of the Afro-Arabian rifts (Prodehl et al., 1997) shows that upper mantle Pn velocities close to 8.0 km/s as found under the Dead Sea (Ginzburg et al., 1981) are more typical of cooler rift flanks than of rifts axes, where slower velocities are usually observed due to anomalous heating. If brittle behavior is then likely to be the dominant deformation mechanism in the lower crust of the basin, strong earthquakes should also nucleate in the lower crust and not only microearthquakes as we report. Unfortunately, due to the long recurrence interval of strong earthquakes in the Dead Sea (Shapira, 1997), several centuries of monitoring might be required before valid statistics would allow a reliable assessment of this likely possibility. As an alternative solution to a brittle lower crust, a model recently developed by Al-Zoubi and ten Brink (2002; ten Brink, 2002) suggests that crustal flow might well be a viable deformation mechanism in the lower crust of the Dead Sea basin. Although this is not what our results mostly suggest, it remains a possibility with a higher heat flow and a lower crust made of wet diabase, or if quartz diorite substitutes for diabase. Around 400°C , a dry diabase would be ductilely too strong and brittle failure would prevail.

If the lower crust of the Dead Sea is cool and brittle, then the upper mantle should also be in a seismogenic state but appears to be aseismic during the 14-year data period. This apparent paradox is in agreement with Maggi et al. (2000), Jackson (2002) and earlier studies (Banda and Cloetingh, 1992; Cloetingh and Burov, 1996) regarding the general scarcity of events in the upper mantle of continents. Cloetingh and Burov (1996) suggest a few possible explanations: stress relaxation due to crust-mantle decoupling, strengthening of the uppermost mantle in downward-bent rift basins and low or inhomogeneous horizontal intraplate stress. Finally, the

mantle lithosphere may also not be so ductilely strong as suggested by laboratory experiments. This last explanation is adopted by Maggi et al. (2000) and Jackson (2002), following the observation (McKenzie and Fairhead, 1997; Maggi et al., 2000) that the effective elastic thickness T_e of the continental lithosphere is usually close to the thickness of the seismogenic crust T_s . Consequently, the mantle would have no significant strength. In the case of the Baikal rift, an effective elastic thickness T_e larger than the total crustal thickness leads Déverchère et al. (2001) to a different conclusion. The authors argue that on the contrary, the high strength of the mantle is the reason explaining the scarcity of earthquakes below the Moho. The upper mantle would not be aseismic but nucleation of earthquakes would be difficult, leading to long recurrence intervals. In the case of the Dead Sea basin, we do not see decoupling between the crust and the mantle as a likely mechanism explaining the absence of earthquakes in the upper mantle, our results pointing towards a mechanically coupled crust-mantle. Unfortunately, this is the only data-driven conclusion we can reach so far about this difficult and controversial subject.

Earthquakes between Aqaba-Elat and the Sea of Galilee were also relocated with the same requirements and velocity model used to derive the distribution of depths in the Dead Sea basin. According to the distribution of focal depths (Figure 2.2), the Dead Sea Transform is not accompanied by any crustal thermal anomaly between the Dead Sea basin and the Sea of Galilee. In this region, deep-crustal seismicity is not only observed along the Dead Sea Transform but also along the Carmel fault zone off the Transform. Southward, the deepest earthquakes quickly become shallower, and barely reach the lower crust in the region of Aqaba-Elat. With appropriate adjustments in rheological models, the maximum depth of approximately 20 km in this region is consistent with the surface heat flow of 65 mWm^{-2} measured in the northern part of the Gulf of Aqaba-Elat (Ben-Avraham and Von Herzen, 1987). The thinning of the seismogenic zone in the region of Aqaba-Elat is probably related more to the increase in the regional heat flow toward the Red Sea (Ben-Avraham and Von Herzen, 1987) than to a heat anomaly directly associated with the Dead Sea Transform itself.

ACKNOWLEDGMENTS

We thank S. Wdowinski for his help during the early stage of rheological modeling. We benefited from stimulating discussions with M. Steckler and S. Feinstein. We are also thankful to Uri ten Brink, Sierd Cloetingh, an anonymous reviewer, and the editor Scott King for their in-depth criticism and advice. F.A. was supported by a Ph.D. grant from the Ministry of Energy and Infrastructure of Israel.

Chapter 3

Methods of a new picking system: MannekenPix

The ability to perform accurate automatic phase picking on a broad range of data remains a serious challenge facing the seismological community. The difficulties involved are somewhat masked by the fact that picking is routinely done just by visual inspection of seismograms. The best picking system widely recognized as such remains the human analyst. But analysts are slow, and sometimes they scamp work due to the boredom induced by the highly repetitive nature of onset picking, especially during periods of intense seismic activity. Analysts also frequently introduce systematic errors due to inadequate working procedures. For instance, such errors can result from the interaction between the picking process and the location results. If the location algorithm does not include distance and residual weighting, chances are that analysts sometimes move accurate picks closer to the predicted arrival in order to reduce the r.m.s. residual time of the computed source locations. As an alternative or complementary mispractice, weights are unduly downgraded at stations affected by high residual arrival times. Sometimes, analysts do not zoom into seismograms (Douglas et al., 1997) to clearly resolve individual samples because of time constraints. Time constraints can also imply that a great deal of seismograms remains unpicked, especially in aftershock sequences and in swarms. Part of the data is then missing for detailed studies. When the picking task is shared among several analysts, inconsistencies appear unavoidable. Even a single analyst is likely to introduce them. It is sufficient to completely repick a few times the same data set to notice many differences in arrival times and weights. If the same exercise is performed annually, then the learning progress of the analyst over time also reveals itself as a source of inconsistencies. The human learning capability, so desirable in itself, comes unfortunately at the price of inconsistencies in the picking and weighting of phases. When several analysts share the picking task as is often the case, their level of expertise is usually uneven, and this makes inconsistencies unavoidable at all times. The highest level of inconsistency is probably reached when data from several networks are merged. In that case,

heterogeneous picking methods, tools and policies add an ultimate layer of inconsistency to the broader data set newly gathered. Inconsistencies can remain unnoticed when the main use of the data is to locate events independently from each other, all phases of any particular event being usually picked by a single analyst. When picking data from different analysts are blended together in studies like joint hypocentral determination and travel-time tomography, inconsistencies appear more clearly. Picking inconsistencies become more apparent as a consequence of the high interaction between all the data in these joint studies. They can severely jeopardize the meaning of the results. The only way to reduce inconsistencies for joint studies is usually to completely revise the picking and the weighting of arrival times.

3.1 AUTOMATIC PICKING

After about thirty years of uncertain progress, a diffuse feeling is that automatic picking still belongs to seismological gadgetry. But the need for high-quality automatic picking is real and quickly growing. Detailed studies are usually performed years after the actual recordings, involving sometimes millions of seismograms. As most of these studies require carefully picked and weighted arrival times, other ways than manually re-picking huge amounts of data are highly desirable. With the abundance of seismograms contaminated by poor practices, automated alternatives are even of vital importance for the feasibility of a growing number of modern seismic applications.

There are basically two main approaches to automated picking. A first way is to pick each seismogram independently from the others and one event at a time. Single-trace picking can be seen as an extension of the detection process, and some methods work in near-real time. It is also the first way analysts usually try during the routine task of picking seismograms. Traditional methods quantify some attributes of waveforms like amplitude, frequency or polarization, and apply their detection and picking algorithms on these attributes or on a smoothed combination of them. Among the wide variety of traditional methods for picking first arrival P-waves of local and regional events, Allen (1978, 1982) designed a STA/LTA algorithm applied on an envelope function sensitive to both amplitude and frequency of seismograms. Despite its age, Allen's picking system is still part of Earthworm (Johnson et al., 1994) and Sac2000 (Goldstein et al., 1999). Baer and Kradolfer (1987) derived their picking engine from Allen's work but they use the square of a modified envelope to generate a characteristic function whose value is tested against an adjustable threshold. The Baer-Kradolfer picker is integrated into Pitsa (Scherbaum and Johnson, 1992). Autoregressive methods (Morita and Hamaguchi, 1984; Takanami and

Kitagawa, 1988; Kushnir et al., 1990; GSE/Japan/40, 1992; Takanami and Kitagawa, 2003) work also well for picking local events, but they can also be used at regional and teleseismic distances (Leonard and Kennett, 1999; Sleeman and van Eck, 1999). The Cusum algorithm (Basseville and Nikiforov, 1993) appears as an attractive alternative to autoregressive methods at regional distances (Der and Shumway, 1999; Der et al., 2000). Klumpen and Joswig (1993) apply pattern recognition on polarization images for P- and S-onset picking of three-component local data. Non-traditional methods like neural networks can sometimes work directly on seismograms (Dai and MacBeth, 1995, 1997), avoiding the need to compute attributes or characteristic functions.

A second approach to automatic picking is to work on several seismograms at once, exploiting the similarity of waveforms from nearby events. This multi-trace approach derives basically from seismic refraction methods (Peraldi and Clement, 1972) used to compute static corrections of seismic reflection data. In seismology, seismograms are typically organized in common station gathers (Dodge et al., 1995; Shearer, 1997). A good account of the methodology and its historical progress is provided by Aster and Rowe (2000). Joswig and Schulte-Theis (1993) extended the application of master-event correlation by using dynamic waveform matching for weak local events occurring in clusters.

The single-trace approach appears today to be more versatile than the multi-trace approach in the sense that it is not based on the restrictive criterion of waveforms similarity. A multi-trace approach is inherently better adapted to relatively restricted volume studies like those where clusters usually occur. Aster and Rowe (2000) suggest however that further developments in adaptive filtering and event clustering (Rowe et al., 2002) might significantly extend the range of data sets that can benefit from cross-correlation picking techniques.

In order to perform appropriately on a wide range of data, I believe that picking algorithms should be integrated into complete processing sequences consistent with the requirements of target applications, where the picking results become the input data. Most picking algorithms need the synergetic support of these complementary components before they can be turned into robust production tools. Despite the impressive number of picking methods reported, the need for complementary components has probably not received enough recognition in the literature. Examples of such components are high-fidelity filters of seismograms, methods for estimating time uncertainties associated with picked phases, and how the deconvolution by the instrumental response could increase the accuracy of phase picking. Estimating picking uncertainties appears to me as the second most important function of an automated system. In order to be meaningful, every physical measurement needs an assessment about its own uncertainty. A valid phase pick

datum should then be understood as made of two quantities: first, the arrival time and second, the uncertainty, both expressed in time units. Many picking algorithms do not tackle the uncertainty estimation at all, or provide qualitative weights poorly tied to time uncertainties.

I think also that the literature has perhaps not done enough to convince potential users that automatic picking can really solve some of their problems. Papers usually focus so much on picking methods that a clear profile of the underlying data is often omitted. Potential users are then unable to see the relation between the data from which the picking algorithms derive and their own data and picking needs. With very few exceptions (Baer and Kradolfer, 1987; Sleeman and van Eck, 1999; Leonard, 2000), comparative studies where several algorithms are applied on a common data set are also missing, making it difficult to know which algorithms perform best, and under which conditions. The lack of comparative studies means also that good results obtained on one data set can be of very little meaning for other sets of data. Some tables summarizing the results of various methods have well been made. Besides classifying the methods involved, these tables do not allow meaningful comparisons as each method is specifically tied to its own data set. On the other hand, it is also often not clear what kind of benefits can be expected from automated picking and at what price. Finally, most users are not willing to perform lengthy tests and related programming efforts for an uncertain result. Users want turn-key solutions that work and these are very difficult to find. Automatic picking capabilities are well integrated into Sac2000 (Goldstein et al., 1999), Pitsa (Scherbaum and Johnson, 1992) and Seisan (Havskov and Ottemöller, 1999), but these packages do not specialize in automatic picking. As such, they also require too much development and tuning efforts for most users before a true production mode can be reached.

3.2 MANNEKENPIX

In order to automatically pick first arrival P-waves on single-component records, I developed a program called MannekenPix. The main goal of the program is to provide picking data close in quality and quantity to those of a good human analyst. Because MannekenPix was originally developed to pick local events in the Dead Sea region where a great variety of waveforms is observed, it follows the single-trace approach. This approach is of interest to the seismological community because of its inherent versatility making it virtually applicable to a broad range of data. MannekenPix has provided good results for local earthquakes from the Dead Sea region (see Chapter 4), and for local - regional earthquakes from Italy (see Chapter 5). These results indicate that the program is probably adaptive enough to satisfy the stringent requirements of

local and regional earthquake tomography for other data sets as well. From version 1.7, MannekenPix is also able to tackle teleseisms but this capability has not yet been tested on a sufficiently large data set.

The program works in guided mode. It searches for a valid arrival around the time of a manual pick or around the time of a predicted value. This is a common setup for high-quality seismological studies where data are selected from a catalogue of routinely located events spanning several years. The presence of more than one event per record is also not uncommon and makes it often desirable to guide the system in search of an arrival of interest. If no event location is available to provide at least a predicted value, the program can also search around the trigger time of a network detection algorithm. MannekenPix is a turn-key solution that does not require much tuning from an experienced user before a true production mode can be reached.

The first benefit that can be expected from MannekenPix is picking accuracy. Most automatic picks are comparable to those of a good human analyst. But as said earlier, the quality of the picking is not only a function of the picking accuracy, it also depends on how well the picking uncertainty is evaluated. MannekenPix includes a weighting mechanism rigorously calibrated on reference picks and weights provided by the user. This calibration is best described as a learning process by example. On unseen data, the weighting engine applies the knowledge gained during the calibration to mark every picked arrival with a quality label. The weighting mechanism of MannekenPix attempts to predict the same uncertainties as those that would be estimated by the user. As such, it acts as an extension of the manual approach and not as a totally independent method possibly in conflict with the legacy. The greatest benefit of the program, an improved picking set, results however from its ability to avoid the mispractices and inconsistencies accumulated by the analysts over the years.

As a side-effect of its own consistency, the program misses some arrivals from seismograms that were picked manually. This can be a problem for events with very few phases because these events might then become un-locatable. However, as tomographic studies usually discard events with very few phases, this should be an affordable price to pay for studies involving great amounts of data. The total number of automatically picked phases can be comparable to the number of phases routinely picked because MannekenPix is usually able to compensate the missed arrivals by picking some decent seismograms left unpicked by the analysts. If many seismograms are routinely unpicked, then the number of seismograms picked by MannekenPix can largely exceed the number of routinely picked seismograms.

In order to increase the signal-to-noise ratio of the P-wavetrain before picking, the seismograms are filtered in the first step of MannekenPix by a high-fidelity Wiener filter. The picking is performed in the second step by the robust and versatile Baer-Kradolfer (Baer and Kradolfer, 1987) picking engine. If a valid pick has been found, a variable delay correction is applied in the third step in order to reduce the inherent delay of the Baer-Kradolfer algorithm arrivals. In the fourth step, the weighting engine provides a statistical estimate of the picking uncertainty. For each data set, the weighting engine has to be calibrated first by a multiple discriminant analysis performed on user-supplied reference picks and weights.

3.3 THE FOUR STEPS OF MANNEKENPIX

The four steps of MannekenPix (Table 3.1) are the foundations of the program, and the methods underlying these steps are reviewed in this section.

Table 3.1 The four steps of MannekenPix

	Processing step	Theory	Implementation
Pre-picking	1. Wiener Filter	Statistical Signal Processing	F.A.
Picking	2. Automatic Picking	Univariate Statistics Signal Processing	Baer-Kradolfer
Post-picking	3. Delay Corrections	Univariate Statistics	F.A.
	4. Weighting	Multivariate Statistics	F.A.

I assume that one of the ultimate goals of automatic picking is to achieve the highest accuracy possible. However, even the highest picking accuracy cannot prevent time base errors to directly affect the total uncertainty on picked phases. Maintaining at all times an accurate time base signal across a network of stations can be difficult, even with GPS timing technology. In a recent LET study of the Ionian channel region in Greece, great efforts were required (Hasslinger, 1998) to reduce time base uncertainties to $\pm 10\text{-}20$ ms (with data sampled every 4 or 10 ms) on radio-

transmitted and GPS time signals of a dense temporary network. The total uncertainty affecting picking arrivals has also to take into account these time base uncertainties, but this can be done separately from the estimation of the picking uncertainties.

3.3.1 WIENER FILTER

Linear least-squares optimal filtering of stationary time series

Following Aki and Richards (1980), signal can be defined as the desired part of the data while noise is the remaining and unwanted part. Signal and noise have thus a relative meaning as they depend on the object of interest. The signal-to-noise ratio is not only affected by the level of ambient ground noise it is also affected by the whole recording process, in which the instrumental response plays an important role.

The attenuation of noise through selective frequency filtering has been widely and often wildly used. Some raw filtering approaches are sometimes described as being optimal. It is improbable that these raw approaches could be optimal in a mathematical sense. I do not mean that raw filters are not useful, but rather that they cannot be optimal if they are not explicitly based on a theory of optimal filters.

As part of the manual picking process, analysts commonly filter seismograms. But manually adjusting corner frequencies and slopes of a filter is time-consuming, and a causal filter can induce further delays in the visible onsets. On the other hand, a noncausal filter can create precursory side-lobes if the corner frequencies and the slopes of the transition bands are not or cannot be perfectly adjusted manually. Another common setup is to apply one common filter to a complete data set. This can improve most seismograms if the filtering effect is mild but in that case, the improvements might also be insufficient. If the filtering effect is stronger, then delays or precursory side-lobes will unavoidably appear on some seismograms, and significant signal attenuation can lead to late picking and polarity errors. As a result, frequency filtering is sometimes simply discarded because the problems it introduces appear to outweigh the benefits. Fortunately, optimal filtering methods exist.

The theoretical foundations of continuous time linear least-squares filters were laid by Norbert Wiener in 1942 in a classified monograph. The work was published later in *Interpolation, Extrapolation and Smoothing of Time Series* (Wiener, 1949). To avoid the transcendental

analysis of continuous time, N. Levinson worked out a discrete-time formulation of Wiener's theory that opened computational possibilities (Levinson, 1947a, 1947b). Today, Wiener filters play an important role in a wide range of applications such as noise reduction, prediction, deconvolution, channel equalization and system identification.

In the frequency domain, a filter output $\hat{S}(f)$ is the product of the filter frequency response $W(f)$ and the total input frequency response $T(f)$:

$$\hat{S}(f) = W(f)T(f). \quad (3.1)$$

The filter error $E(f)$ is defined as the difference between a desired signal $S(f)$ and the filter output $\hat{S}(f)$:

$$E(f) = S(f) - \hat{S}(f). \quad (3.2)$$

The mean-square error ξ of the filter is then

$$\xi = \mathbf{E}\left\{ |E(f)|^2 \right\} \quad (3.3)$$

with \mathbf{E} being the expectation.

A Wiener filter is a linear operator designed to minimize the mean-square error (3.3) between the filter output and a desired signal. In their basic form, Wiener filters assume that signal and noise are wide-sense stationary (WSS) random processes. If signal and noise are Gaussian time series, then the Wiener filter is an optimal estimator and it is also optimal among all non-linear estimators. This property is not true for non-Gaussian signal and noise.

It can be shown (Hayes, 1996; Vaseghi, 2000) that the minimum mean-square error Wiener filter in the frequency domain is given by

$$W(f) = \frac{P_{S,T}(f)}{P_{T,T}(f)} \quad (3.4)$$

where $P_{S,T}(f)$ is the cross-power spectrum between the signal $S(f)$ and the total input $T(f)$, and $P_{T,T}(f)$ is the power spectrum of the total input $T(f)$.

In the case of a signal $S(f)$ buried in additive noise $N(f)$, the total input $T(f)$ is then

$$T(f) = S(f) + N(f). \quad (3.5)$$

Assuming that the signal $S(f)$ and the noise $N(f)$ are uncorrelated, we can write

$$P_{S,T}(f) = P_{S,S}(f) \quad (3.6)$$

and

$$P_{T,T}(f) = P_{S,S}(f) + P_{N,N}(f) \quad (3.7)$$

where $P_{S,S}(f)$ is the power spectrum of the signal and $P_{N,N}(f)$ is the power spectrum of the noise. The Wiener filter in the case of uncorrelated additive noise is then given by

$$W(f) = \frac{P_{S,S}(f)}{P_{T,T}(f)} = \frac{P_{S,S}(f)}{P_{S,S}(f) + P_{N,N}(f)}. \quad (3.8)$$

From equation (3.8), the Wiener filter for noise reduction is often described as a $S^2/(S^2 + N^2)$ filter. The amplitude of $W(f)$ will be close to 1 for frequencies where the noise is negligible, and close to 0 where the noise is dominant. The intermediate values given by (3.8) are the optimal way of filtering the noise (maximize the SNR) between the two extremes. As a consequence of the linear framework, phase angles of signal and noise are not used (Bode and Shannon, 1950) in Wiener filters.

It can be further shown (Press et al., 1986; Hayes, 1996) that equation (3.8) remains valid if the desired signal is not the recorded signal $S(f)$, but the true original signal $S_0(f)$ unaffected by the instrumental response. In that case, the filtering problem can be solved first for the estimated recorded signal $\hat{S}(f)$, leading from equations (3.1) and (3.8) to

$$\hat{S}(f) = \frac{P_{S,S}(f)}{P_{S,S}(f) + P_{N,N}(f)} T(f). \quad (3.9)$$

The estimate $\hat{S}_0(f)$ of the true original signal $S_0(f)$ is then

$$\hat{S}_0(f) = \frac{\hat{S}(f)}{R(f)} \quad (3.10)$$

where $R(f)$ is the impulse response of the recording system. The full Wiener filter is then made of frequency filtering (3.9) followed by deconvolution (3.10). The frequency filtering regularizes the deconvolution by suppressing the noise at frequencies where it dominates the signal, and it only requires the power spectrum of the recorded signal $S(f)$, not the power spectrum of the true original signal $S_0(f)$.

In seismology, removing the response of the recording system is usually performed for the determination of ground motion amplitudes, and not for picking purposes. However, deconvolving the instrumental response is part of the full Wiener filter and it might become an advanced tool in search for the most accurate onsets of seismic phases, whether picked manually or automatically. The high-frequency content of short-period seismograms might however

suggest that removing the response of the recording system is not worth the burden compared to the potential accuracy recoverable. This might be true for the routine work but I think that deconvolution should not be overlooked for high-accuracy studies.

The main problem with the recording system is that it usually distorts and lengthens significantly the true ground motion, whether it be signal or noise. Such a systematic distortion does not only mean that recorded waveforms do not accurately reflect the true shape of signals, it also implies a reduction of the signal-to-noise ratio. The seismic recording system acts as an out-of-focus lens through which everything appears blurred, even in the absence of additive noise. On most seismograms, the amplitude of the first samples of seismic phases is so low that true onsets are nearly always obscured by some residual noise. This contrasts with the sharp initial rise suggested by source models, especially when the data are recorded by short-period seismometers mainly sensitive to velocity. This problem leads some analysts to systematically and often arbitrarily pick earlier than the onsets they really see. Under these circumstances, accurately picking seismic phases appears to be rather illusory. A synthetic example will show that the instrumental response can explain a great deal of the blur observed, even for noiseless seismograms. Deconvolution would then be the way to refocus the smeared signals. If widely applicable to real-world data within the scope of phase picking, deconvolution could significantly increase the accuracy of single-trace picking and reduce related uncertainties.

Stochastic and data-adaptive Wiener filters

In their initial form, Wiener filters are based on ensemble averages as in equation (3.3), and one filter is applied to all realizations of a given process. This approach is called *stochastic* by Berkhout and Zaanen (1976). Models for the noise and for the signal are required. As a second approach and under the assumption that ensemble averages can be approximated by time averages, Wiener filters can also be *data-adaptive* or *deterministic*. They can then be tailored to specific realizations where the statistical properties of the signal and noise are directly estimated from the data. While background noise is often almost stationary over short segments of time, seismic signals are transients. It is clear then that strict stationarity of noise and signal cannot be expected from seismograms. One way to apply Wiener filters is to restrict their use or validity to data segments where signal and noise are almost stationary. In the framework of guided phase picking where an arrival is constrained to occur within some time window, segmented stationarity is usually applicable.

Among the few Wiener filters for noise reduction described in the seismological literature, Douglas (1997) follows the stochastic approach and computes the filter in the time domain according to Franklin's method (1970). In the presence of errors, the linear inverse problem can be written as

$$\mathbf{A}\mathbf{s} + \mathbf{n} = \mathbf{t} \quad (3.11)$$

where \mathbf{s} is the unknown, \mathbf{t} the observation, \mathbf{n} the error vector, and \mathbf{A} is the coefficient matrix relating \mathbf{s} and \mathbf{t} . In the case where \mathbf{t} is a sampled times series containing a signal \mathbf{s} buried in additive uncorrelated noise \mathbf{n} , \mathbf{A} is the identity matrix and the Wiener filter \mathbf{w} is given in the time domain by

$$\mathbf{w} = \mathbf{R}_{s,s} [\mathbf{R}_{s,s} + \mathbf{R}_{n,n}]^{-1} \quad (3.12)$$

where $\mathbf{R}_{s,s}$ is the autocorrelation matrix of the signal and $\mathbf{R}_{n,n}$ is the autocorrelation matrix of the noise. The similarity between (3.8) and (3.12) results from the Wiener-Kintchine theorem (Wiener, 1930; Kintchine, 1934) stating that autocorrelation and power spectrum are Fourier transforms of each other. The estimate $\hat{\mathbf{s}}$ of the signal \mathbf{s} is then given by

$$\hat{\mathbf{s}} = \mathbf{R}_{s,s} [\mathbf{R}_{s,s} + \mathbf{R}_{n,n}]^{-1} \mathbf{t} . \quad (3.13)$$

In order to filter only the low-frequency part of the noise, Douglas uses a tapered cosine wave of the predominant noise period as model for the noise, and the impulse response of the recording system as model for the signal. The author expresses his preference for filters derived from models over data-adaptive filters. First, a great number of data points are required to obtain detailed and reliable autocorrelations of noise segments, and stationarity becomes an issue for autocorrelations derived from long data segments. When only a small number of points is used, the resulting filter has a smooth response and the detailed features of the noise sample are lost. Data-adaptive filters thus add complexity for little advantage. Further, the smoothness of filters derived from models helps in reducing the number of artifacts on the filtered seismograms.

The Wiener filter of MannekenPix is however data-adaptive, signal and noise power spectra in (3.8) being automatically estimated from the data for each seismogram. A first window located ahead of the a priori arrival time samples the noise, while part of the noisy P-wavetrain is sampled by a second window located after the arrival. Assuming as before that the signal and the noise are uncorrelated, the power spectrum $P_{SN,SN}(f)$ of the noisy signal $SN(f)$ is then

$$P_{SN,SN}(f) \approx P_{S,S}(f) + P_{N,N}(f) \quad (3.14)$$

and approximates thus the denominator of the Wiener filter (3.8). The numerator of the filter (the power spectrum $P_{S,S}(f)$ of the signal) can be estimated by subtracting the power spectrum $P_{N,N}(f)$ of the noise from (3.14). The Wiener filter of MannekenPix can then be written as

$$\boxed{W(f) = \frac{P_{S,S}(f)}{P_{S,S}(f) + P_{N,N}(f)} \approx \frac{P_{SN,SN}(f) - P_{N,N}(f)}{P_{SN,SN}(f)}} \quad (3.15)$$

The main difference between Douglas stochastic filter and the data-adaptive filter of MannekenPix is that the scope of the later is broader compared to the former. Douglas filter attenuates only the low-frequency noise and considers the signal to be highly impulsive in all cases. While these sensible assumptions can help to reduce artifacts, they might also be too restrictive. Recorded seismic signals are rarely impulsive enough to fill the bandwidth of the recording system, and high-pass filtering is then not always sufficient. By adapting itself to the frequency content of each seismogram, the Wiener filter of MannekenPix acts as a high-pass, a low-pass, a band-pass, a notch filter, or any meaningful combination of these primitives. Picking routines such as the Baer-Kradolfer algorithm highly amplify both amplitude and frequency changes of seismograms. Any frequency dominated by noise should then be attenuated as much as possible, whether it be noise located in a single band of frequencies or in multiple bands.

Douglas objection against such detailed filters results from the apparent impossibility to satisfy both practical stationarity (as with short data segments) and high resolution (as with long data segments). But solutions to this old problem are already known for more than thirty years and they cannot be ignored today.

Power spectrum estimation

By virtue of the Wiener-Kintchine theorem (Wiener, 1930; Kintchine, 1934) stating that autocorrelation and power spectrum are a Fourier transform pair, estimating the power spectrum is equivalent in the frequency domain to estimating the autocorrelation in the time domain. In theory, all that must be done then is to compute the autocorrelation and take its Fourier transform. However, there are two difficulties with this direct approach. First, practical stationarity of both the noise and the transient seismic signal requires the use of short data segments but detailed and reliable spectra are difficult to derive from short segments. The second difficulty is that estimating power spectra of real-world data is always done in a noisy environment directly affecting the results.

The numerous methods for power spectrum estimation are generally categorized into at least two main classes. The first class includes *classical* or *nonparametric* methods, in which the power spectrum is directly estimated from the data. Classical methods differ mainly in the way they window the data or the autocorrelation before taking a Fourier transform. They are also sometimes referred to as linear methods because they only involve linear operations, and their window functions are independent of the data. The second class includes *nonclassical* or *parametric* methods, in which the data are assumed to be the output of a linear system driven by white noise. Parametric methods are based on a model for the process generating the data. They are said to be nonlinear because their window functions are data-adaptive. The maximum entropy method (MEM) appears as a logical first step in the world of parametric methods since its main properties for 1-D spectral analysis (MEM1) are rather well established both theoretically and experimentally (Wu, 1997), and numerous implementations of the method are readily available.

Classical methods are more general than parametric methods since they do not incorporate information about the process that generated the data. In some applications however, this may be an important limitation. In speech processing, an acoustic model for the vocal tract imposes an autoregressive model on the speech waveform (Rabiner and Schafer, 1978). For short intervals of time over which the speech waveform is approximately stationary, a parametric approach based on an autoregressive model could then provide a more accurate estimate of the power spectrum than a nonparametric method.

Classical methods of power spectrum estimation

The existence of numerous classical methods for spectral estimation results mainly from attempts to overcome the inherent limitations of the *Periodogram*, a method first introduced by Schuster (1898) in his study of periodicities in sunspot numbers. For a discrete signal $x(n)$ of length N , the periodogram $\hat{P}_{x,x}(f)$ is proportional to the squared magnitude of the discrete Fourier transform (DFT) of $x(n)$ and is given (e.g. Hayes, 1996) by

$$\hat{P}_{x,x}(f) = \frac{1}{N} |X(f)|^2 \quad (3.16)$$

where $X(f)$ is the N -point discrete Fourier transform (DFT) of $x(n)$

$$X(f) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi n f} \quad (3.17)$$

With respect to the autocorrelation, a finite segment of data behaves as an infinite series multiplied by a Bartlett (triangular) window. Multiplication in the time domain by the Bartlett window corresponds to a convolution in the frequency domain by the Fourier transform of the Bartlett window. The periodogram is a biased estimate of the power spectrum and its expectation is given by

$$E\{\hat{P}_{X,X}(f)\} = P_{X,X}(f) * W_B(f) \quad (3.18)$$

where $P_{X,X}(f)$ is the true power spectrum, “*” denotes convolution, and $W_B(f)$ is the Fourier transform of the Bartlett window, converging to an impulse as N goes to infinity. The periodogram is then asymptotically unbiased

$$\lim_{N \rightarrow \infty} E\{\hat{P}_{X,X}(f)\} = P_{X,X}(f). \quad (3.19)$$

For white Gaussian noise, the variance of the periodogram is proportional to the square of the power spectrum

$$Var\{\hat{P}_{X,X}(f)\} \propto P_{X,X}^2(f) \quad (3.20)$$

and does not go to zero as N goes to infinity. The periodogram is thus not a consistent estimate of the power spectrum.

Figure 3.1 displays the periodogram power spectral density (PSD) of a noiseless harmonic signal, computed with function “periodogram” of Matlab Signal Processing Toolbox. This periodogram derives from a long data segment encompassing more than 50 periods of the monochromatic signal. The same signal will serve to illustrate the main properties of other widely used methods of estimating power spectral densities. A one-sided PSD means that the integrated power spectral density over only positive frequencies equals the total power (mean-square amplitude) of the data segment. In order to increase the number of points in the frequency domain, the data segment was zero padded up to a total of 2,048 points by the “FFT” function of Matlab. There are then $(2,048 / 2) + 1 = 1,025$ frequencies on the positive axis between 0 Hz and the Nyquist frequency (50 Hz), and the frequency sampling period is 0.0488 Hz. Due to the Bartlett window inherent of the periodogram, significant power is transferred (spectral leakage) to the whole range of frequencies. In addition, the smoothing introduced by windowing broadens spectral features and hence reduces resolution. In the example, the main lobe at 20.5 Hz is 0.55 Hz wide at the 10-percent level of the peak value.

Even with such a long data segment as the one of the example, the periodogram suffers thus a rather poor resolution and a high spectral leakage. It does not appear well suited for analyzing closely-spaced sinusoids (due to low resolution) and continuous spectra (due to high leakage). In

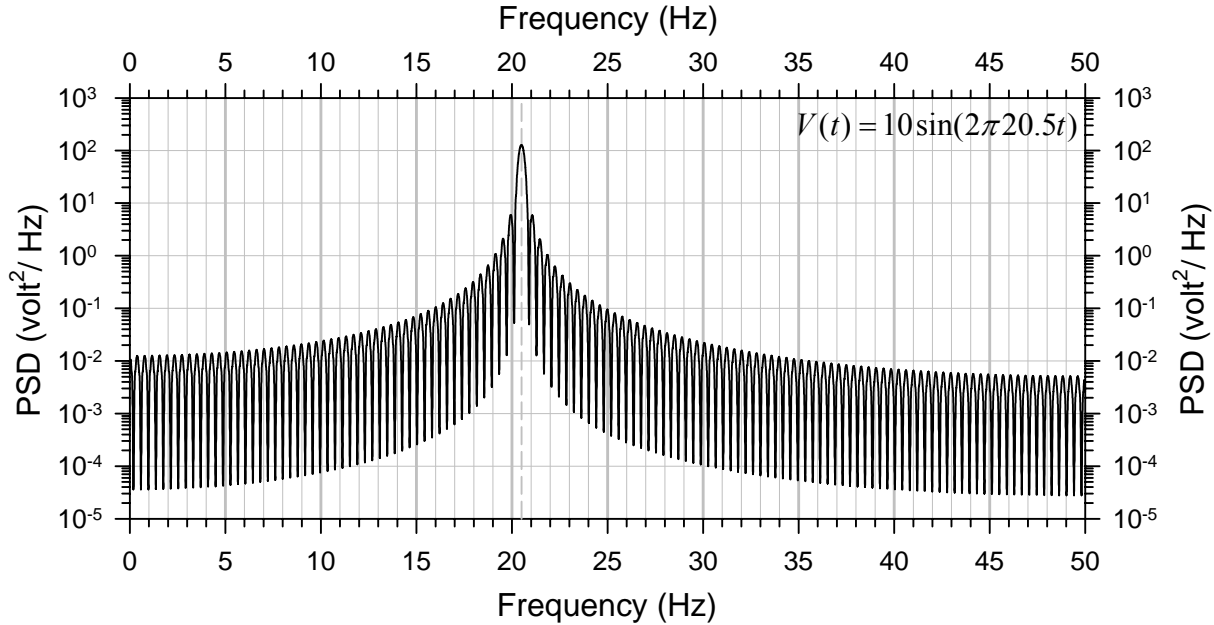


Figure 3.1 One-sided *periodogram* of a 20.5 Hz harmonic potential difference $V(t)$ of 10 volts peak amplitude. 256-point data segment, 0.010 sec sampling period, and 2,048-point FFT.

addition, bias and resolution further degrade when short data segments are used. This is especially important in our case since we need spectra derived from short windows in order to ensure practical stationarity of the data segments used to estimate the Wiener filter.

For any finite data segment of N samples, the number of samples available for the estimation of the autocorrelation value $\hat{r}_{x,x}(k)$ at lag k , is very low as $|k|$ approaches N . Therefore, the estimate of the autocorrelation at lag k when $|k|$ is close to N , is always unreliable and the variance of the entire autocorrelation estimate increases. In the *Blackman-Tukey method* (1958), also referred to as *periodogram smoothing*, the variance is reduced by applying a (correlation lag) window to the autocorrelation estimate before taking its Fourier transform. What is traded for this reduction in variance, however, is a reduction in resolution since a smaller number of (and now unequally weighted) lags is used to estimate the power spectrum.

Figure 3.2 displays the Blackman-Tukey method PSD of the example signal, computed with Matlab code from Hayes (1996). A 384-point wide Blackman lag window produced a smooth spectrum, but the 20.5 Hz peak is now 1.25 Hz wide at the 10 percent level. In this method, the Blackman-windowed autocorrelation is computed using only the 256 available data points, and this lagged autocorrelation is then zero padded when the FFT is taken.

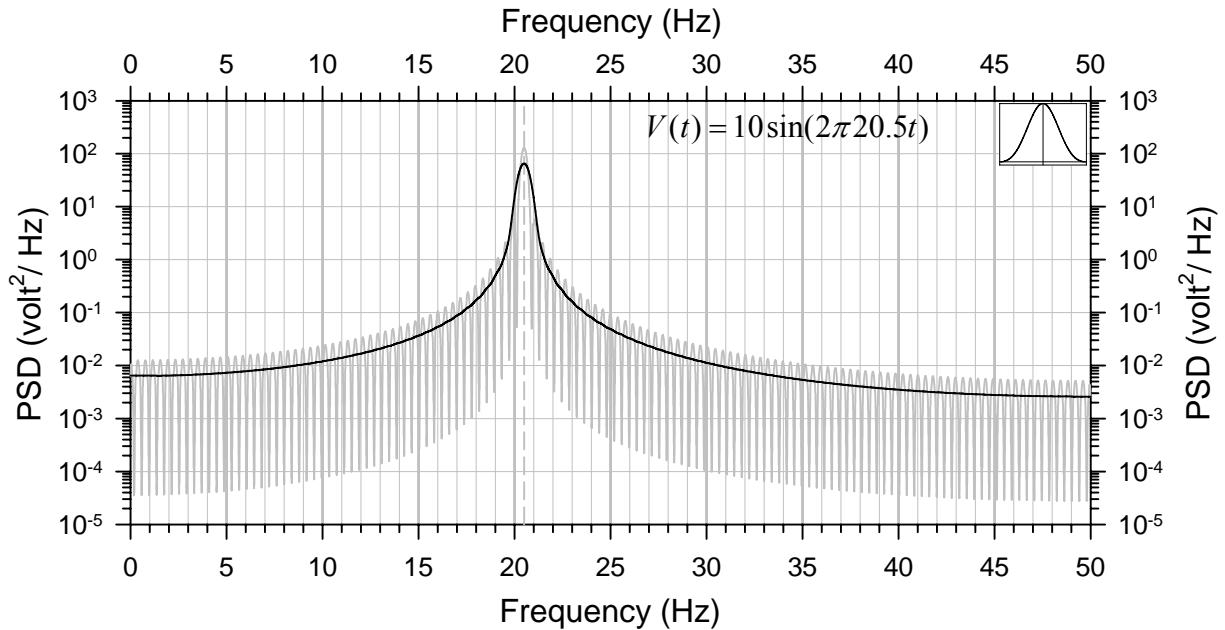


Figure 3.2 One-sided *Blackman-Tukey method* PSD of a 20.5 Hz harmonic potential difference $V(t)$ of 10 volts peak amplitude. 256-point data segment, 0.010 sec sampling period, 2,048-point FFT, and 384-point Blackman window (its shape is in the upper-right corner). The periodogram is shown in gray for comparison.

In Thomson's *Multitaper method* (Thomson, 1982; Park et al., 1987; Percival and Walden, 1993), a small number of leakage-resistant tapers is applied to a single time series from which several spectra are produced. The tapers, known as discrete prolate spheroidal sequences (DPSS), are constructed so that each taper samples the time series in a different manner while optimizing resistance to spectral leakage. The statistical information discarded by the first taper is partially recovered by the next taper, and so on. Averaging over this small ensemble of spectra reduces the variance compared to single-taper methods. In the multitaper method, the time-bandwidth product parameter NW allows to balance the tradeoff between variance and resolution. There are $2NW - 1$ tapers used, so that when NW increases, there are also more estimates of the power spectrum and the variance of the averaged estimate decreases. However, the bandwidth of each taper is also proportional to NW and spectral leakage increases as NW increases.

Figure 3.3 displays the Multitaper method PSD of the example signal, computed with Matlab code from P.J. Huybers¹. The time bandwidth product NW was set at 1.5, resulting in only 2 DPSS tapers. The 20.5 Hz peak is 1.15 Hz wide at the 10 percent level, only slightly narrower

¹<http://web.mit.edu/phuybers/>

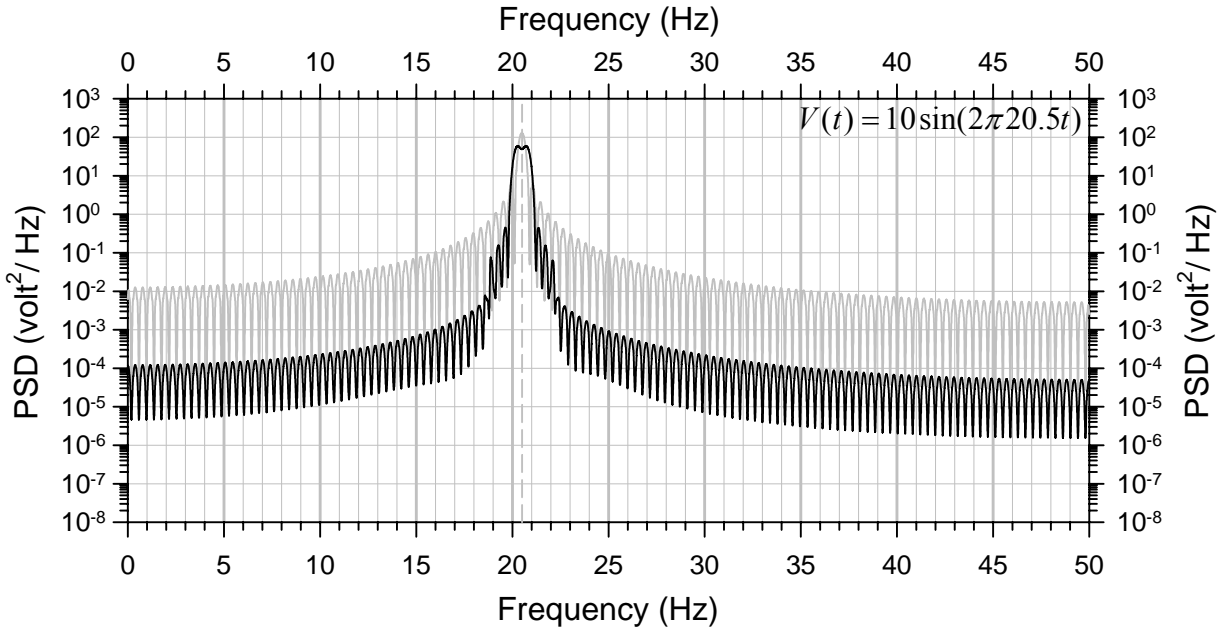


Figure 3.3 One-sided *Multitaper method* PSD of a 20.5 Hz harmonic potential difference $V(t)$ of 10 volts peak amplitude. 256-point data segment, 0.010 sec sampling period, 2,048-point FFT, and $NW = 1.5$ (2 tapers). The periodogram is shown in gray for comparison.

than with the Blackman-Tukey method, but spectral leakage appears significantly reduced compared to previous methods.

The Maximum Entropy Method (MEM)

A serious limitation of classical methods to power spectrum estimation is that, for a data segment of length N , the autocorrelation $r_{x,x}(k)$ can only be estimated for $|k| < N$, and $r_{x,x}(k)$ is assumed to be zero for $|k| \geq N$. Since many signals have autocorrelations that are nonzero for lags $|k| \geq N$, this assumption often results in significant spectral leakage and loss of spectral resolution. This is particularly the case for narrowband signals, for which autocorrelations decay slowly with k . In general, there is an infinite number of autocorrelations that are identical to the estimated autocorrelation of a data segment for lags $|k| < N$. Therefore, constraints must be imposed on the set of possible extrapolations in order to provide a meaningful and unique answer.

Given the autocorrelation $r_{x,x}(k)$ of a WSS process $x(n)$ for $|k| \leq m$, the problem is how to extrapolate $r_{x,x}(k)$ for $|k| > m$. In the frequency domain, an extrapolated power spectrum $P_{X,X}(f)$ can be written as

$$P_{X,X}(f) = \sum_{k=-m}^m r_{x,x}(k) e^{-j2\pi fk} + \sum_{|k|>m} r_e(k) e^{-j2\pi fk} \quad (3.21)$$

where $r_{x,x}(k)$ is the data-derived autocorrelation, and $r_e(k)$ is the extrapolated autocorrelation. First, $P_{X,X}(f)$ needs to be a valid power spectrum, and hence it should be real-valued and nonnegative for all f . An additional constraint, proposed by Burg (1967), is to extrapolate the autocorrelation in such a way as to maximize (Jaynes, 1957a, 1957b) the information entropy (Shannon, 1948) of the process. Since entropy is a measure of randomness, a maximum entropy estimate provides the autocorrelation sequence corresponding to the most random signal whose correlation values for lags $|k| \leq m$ coincide with the measured values. Such an extrapolation places as few constraints as possible or the least amount of structure on the process $x(n)$.

For a Gaussian random process with power spectrum $P_{X,X}(f)$, the entropy rate (e.g. Smylie et al., 1973; Wu, 1997) is given by

$$H = \frac{1}{2} \int_{-1/2}^{1/2} \ln P_{X,X}(f) df. \quad (3.22)$$

For Gaussian processes with a given partial autocorrelation sequence $r_{x,x}(k)$ for $|k| \leq m$, the maximum entropy power spectrum is the power spectrum that maximizes equation (3.22) subject to the constraint that the inverse discrete-time Fourier transform of $P_{X,X}(f)$ equals the partial autocorrelation sequence $r_{x,x}(k)$ for $|k| \leq m$,

$$r_{x,x}(k) = \int_{-1/2}^{1/2} P_{X,X}(f) e^{j2\pi fk} df, \quad |k| \leq m. \quad (3.23)$$

The variational procedure leads to the *Maximum Entropy Method* power spectrum (Burg, 1967, 1975) $\hat{P}_{MEM}(f)$, given for a unit sampling by

$$\hat{P}_{MEM}(f) = \frac{P_{m+1}}{\left| 1 + \sum_{k=1}^m a_{mk} e^{-j2\pi fk} \right|^2} \quad (3.24)$$

where P_{m+1} and the coefficients a_{mk} are given by the matrix equation

$$\begin{bmatrix} r_{x,x}(0) & r_{x,x}(1) & \cdots & r_{x,x}(m) \\ r_{x,x}(1) & r_{x,x}(0) & \cdots & r_{x,x}(m-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_{x,x}(m) & r_{x,x}(m-1) & \cdots & r_{x,x}(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{m1} \\ \vdots \\ a_{mm} \end{bmatrix} = \begin{bmatrix} P_{m+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (3.25)$$

Equation (3.25) is equivalent to the design of a $(m+1)$ -point prediction error filter with mean square error P_{m+1} , and whose first coefficient is set equal to 1. van den Bos (1971) has shown that the maximum entropy method is also equivalent to least-squares fitting of a discrete-time autoregressive (all-pole) model (Yule, 1927) of some order M to the data. An order M autoregressive process $x(t)$ with unit sampling can be written as

$$x(t) = \sum_{k=1}^M a_{mk} x(t-k) + \xi(t) \quad (3.26)$$

where the white noise contribution $\xi(t)$ is usually called the *innovation* of the AR process. The constants a_{mk} in equation (3.26) are the coefficients of the AR process and are solutions to the matrix equation (3.25). The mean square error P_{m+1} of the prediction error filter in equations (3.24) and (3.25) corresponds, for the AR model, to the mean square value $E(\xi^2(t))$ of the innovation.

Several methods can be used to solve the linear matrix equation (3.25) but once the all-pole coefficients a_{mk} have been estimated, each method generates the power spectrum $\hat{P}_{MEM}(f)$ according to equation (3.24).

Burg's algorithm

The autocorrelation matrix $\mathbf{R}_{x,x}$ in equation (3.25) has a *Toeplitz* structure, and the *Levinson-Durbin recursion* (Levinson, 1947a; Durbin, 1960) can be used to solve this equation. The direct solution to equation (3.25), often referred to as the *Yule-Walker method* (Yule, 1927; Walker, 1931), uses traditional sample correlations to estimate the autocorrelation values $r_{x,x}(k)$.

Equation (3.25) of the maximum entropy method assumes however that the autocorrelation lags $r_{x,x}(k)$ are precisely known, but windowing end-effects are unavoidable when the autocorrelation lags are directly evaluated from the data. In order to avoid these end-effects and keep the potential resolution of MEM, Burg developed an algorithm to evaluate the coefficients a_{mk} that does not use the estimated autocorrelation lags $r_{x,x}(k)$ as input. The basis of the algorithm is that the first $(N+1)$ lags of the autocorrelation can be estimated by evaluating the corresponding $(N+1)$ -point prediction error filter. Burg's recursive algorithm (Burg, 1968 and 1975; Andersen, 1974) is based on minimizing the average (forward and backward) prediction error filter output power, starting from a model order equal to 0 and stepwise (Levinson's recursion) increasing the matrix dimension in equation (3.25) until the last coefficient a_{mm} for a

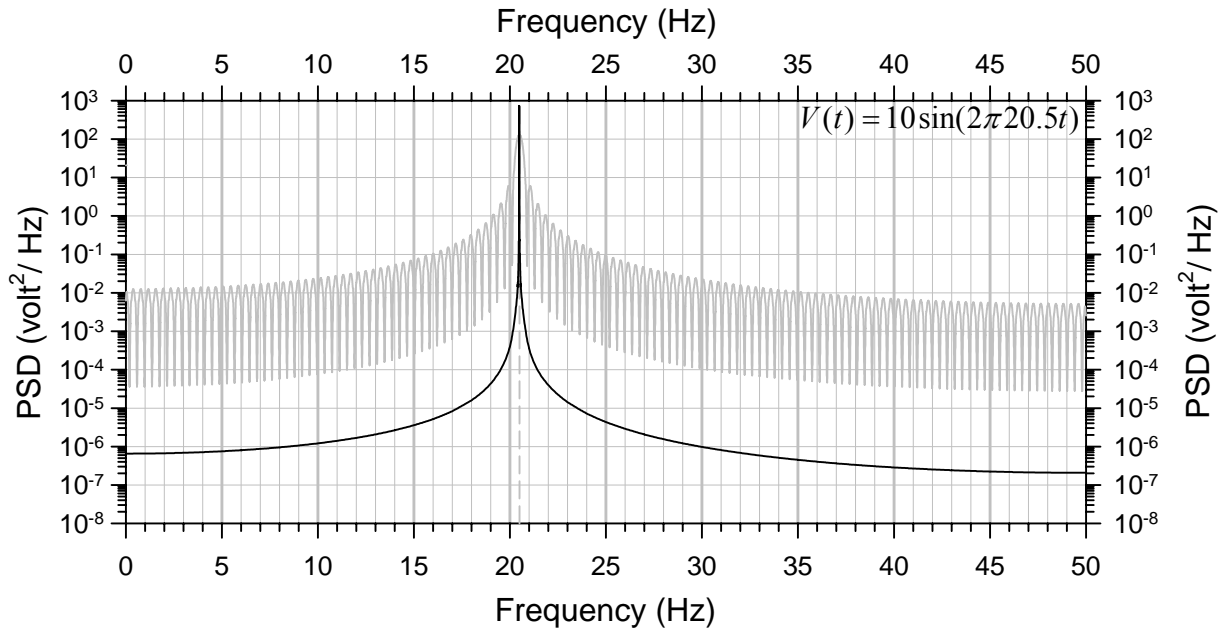


Figure 3.4 One-sided *Maximum Entropy Method* (Burg’s algorithm) PSD of a 20.5 Hz harmonic potential difference $V(t)$ of 10 volts peak amplitude. 256-point data segment, 0.010 sec sampling period, AR order = 2, 10,000 frequencies. The periodogram is shown in gray for comparison.

chosen order M is computed. It is important to note that with Burg’s algorithm, the prediction error filter never runs off the ends of the data sample. The usual assumptions about the time series like zeros before and after the data sample, or periodicity of the data segment, are thus not made. Instead, Burg’s algorithm makes an implicit prediction of the signal beyond the observation window.

Figure 3.4 displays the Maximum Entropy Method PSD of the example signal, computed with function “pburg” (Burg’s algorithm) of Matlab Signal Processing Toolbox. The AR model order M is 2 and the spectrum is evaluated at 10,000 frequencies. For signals composed of very sharp spectral features as in this example, MEM usually requires a higher number of frequencies than classical methods do because a great deal of the power can be concentrated in only a few very closely spaced frequencies. Contrary to classical methods, MEM does not perform zero padding when the desired number of frequencies exceeds the number of data points. In Figure 3.4, the peak is so sharp that there is no sample close to the 10 percent level on both sides of the peak, even when 10,000 frequencies are used. With this number of frequencies, the 20.5 Hz peak is only 0.02 Hz wide at the 0.1 percent level of the peak value. The main properties of this MEM spectrum are thus high sharpness of the peak, and low, smooth leakage.

Selection of the model order in MEM

The duality between MEM and the fitting of an autoregressive model has allowed to apply results from the AR literature to the maximum entropy method. In particular, objective criteria to the selection of the order M in autoregressive models have significantly contributed to MEM. If the model order is too small, the resulting spectrum will be too smooth and will have poor resolution. On the other hand, if the model order is too large, then the spectrum may contain spurious peaks and spectral line splitting may result.

Akaike's Final Prediction Error (Akaike, 1969) chooses the order of the autoregressive model by minimizing the average error power for a one step prediction, considering both the error power due to the unpredictable part of the process (the innovation) and the error power due to inaccuracies in estimating the AR coefficients. With the mean of the data sample removed, the FPE is given by

$$\text{FPE}(M) = \mathcal{E}_M \frac{N + M + 1}{N - M - 1}. \quad (3.27)$$

In FPE, N is the number of points of the observed data segment, M is the AR model order and \mathcal{E}_M is the modeling error power for order M . Akaike's An Information theoretic Criterion (Akaike, 1974) is an extension of this model selection to any maximum likelihood situation. For an AR process with Gaussian errors, and data sample mean removed, the AIC is

$$\text{AIC}(M) = N \ln \mathcal{E}_M + 2(1 + M). \quad (3.28)$$

In both methods, the optimal order M_0 is the model order that minimizes the selected criterion (FPE or AIC). AIC and FPE are asymptotically equivalent and nearly always select the same optimal order M_0 (Jones, 1974).

Jones (1976) has further shown that the FPE criterion, developed originally for the Yule-Walker method of solving equation (3.25), applies also to Burg's algorithm as long as the model order M is not greater than $N/2$. For higher orders, the FPE criterion provides overestimated M_0 values. This is especially important when Burg's algorithm is used (Ulrych and Bishop, 1975) since the variance of the results increases rapidly with this algorithm when the order of the model is overestimated. Jones (1976) has well proposed an empirical correction to FPE for Burg's algorithm when $N/2 < M < N - 1$, but the practice however is generally to avoid model orders greater than $N/2$.

Finally, Landers and Lacoss (1977) have shown that FPE or AIC criteria provide acceptable spectra for harmonic processes with low levels of white noise. For higher noise levels (50%), these criteria gave inconsistent and underestimated model orders. The authors conclude that in general, FPE and AIC should be best seen as approximate indicators of the model order and not as absolute criteria applicable in all cases.

Properties of an MEM spectrum

Because of the high degree of nonlinearity of MEM, its properties are significantly more difficult to ascertain than for classical methods. A full account of the known properties can be found in Wu, 1997. The maximum entropy method has three main properties that make it an important method for the design of an effective Wiener filter. I assume that MEM is implemented according to Burg's algorithm. The first property is that MEM tends to produce sharper spectral features than classical methods do. This is a desirable property since the Wiener filter can then be more tightly shaped around true spectral features. The second property of MEM is the smoothness of spectra and low leakage since no windowing function is applied to the data. This is also an important property since smooth spectra and low leakage are less likely to introduce artifacts than spectra affected by numerous sidelobes. Finally, the most important property is that MEM spectra for short data segments tend to be significantly better than those derived from classical methods. This is particularly important in our case since stationarity requirements of the Wiener filter impose the use of short data segments for spectral estimation.

Figure 3.5 displays an example of power spectral densities derived from a synthetic noise segment of short duration (2 seconds) using the briefly reviewed methods of previous sections. Three narrow beams of noise have been synthesized, with central frequencies located at 0.75, 2.00 and 5.00 Hz. Each beam is made up of three pure sinusoids, the amplitude at each central frequency being twice the amplitude at the two adjacent frequencies within each beam (Table 3.2). In order to make the example more realistic, I added a low level (variance of 4.0×10^{-10} volts) of white Gaussian noise (Figure 3.5b) representing random noise and measurement errors. The resulting waveform (Figure 3.5a) is thus composed of a deterministic component of noise and of a random component of noise. Since only the deterministic part of the noise creates sharp spectral features, it can be seen as the "signal" from the perspective of spectral analysis. The background random noise is then the "noise". The signal-to-noise ratio in this example is equal to 28 dB, a very high value. The power spectra have been computed in Matlab with the same functions that were used in the example of a monochromatic signal. In the case of the MEM

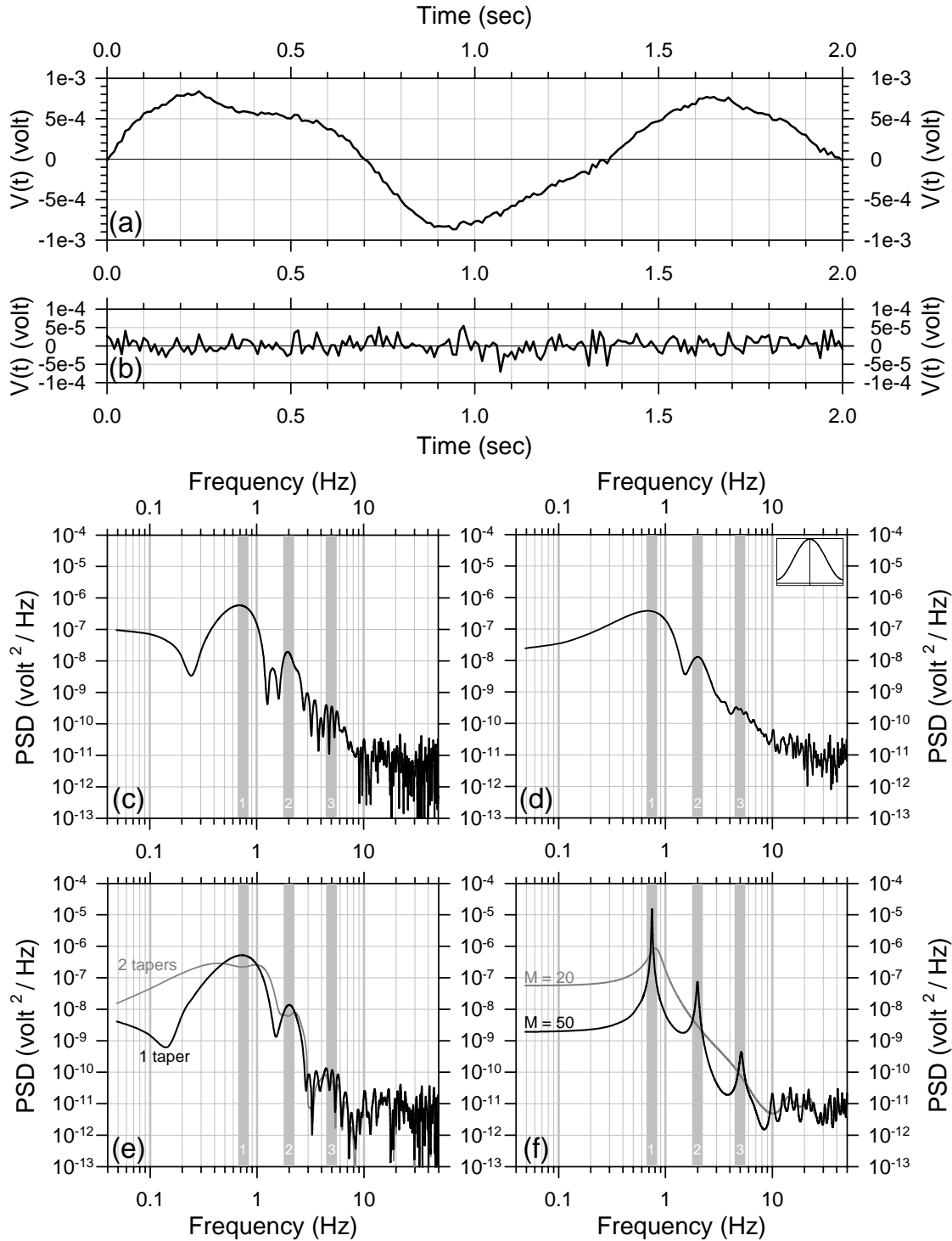


Figure 3.5 One-sided power spectral densities from a synthetic noise segment of short duration. (a) The noise segment, expressed in volts (white Gaussian noise added): 201-point segment, 0.010 sec sampling period. (b) White Gaussian noise of 4.0×10^{-10} volts variance. (c) Periodogram PSD: 8,192-point FFT. (d) Blackman-Tukey PSD: 8,192-point FFT and 201-point Hamming window (its shape is in the upper-right corner). (e) Multitaper PSD: 8,192-point FFT, 1 taper (black) and 2 tapers (dark gray). (f) Maximum entropy (Burg's algorithm) PSD: 8,192 frequencies, AR orders 20 (dark gray) and 50 (black). The three beams of deterministic noise are displayed in gray and labeled in white on the PSD plots.

spectrum, the optimal model order was determined according to the FPE criterion, the modeling error \mathcal{E}_M being a possible output of function “arburg” which computes the AR coefficients in the Signal Processing Toolbox of Matlab.

Table 3.2 Parameters of the deterministic part of the noise in Figure 3.5

Beam #	Peak Amplitude A (volt)	Frequency F (hertz)	Phase ϕ (radian)
1	2×10^{-4}	0.675	0
1	4×10^{-4}	0.750	0
1	2×10^{-4}	0.825	0
2	5×10^{-5}	1.800	$-\pi / 16$
2	1×10^{-4}	2.000	$-\pi / 16$
2	5×10^{-5}	2.200	$-\pi / 16$
3	1×10^{-5}	4.500	$-\pi / 32$
3	2×10^{-5}	5.000	$-\pi / 32$
3	1×10^{-5}	5.500	$-\pi / 32$

Potential difference as a function of time is given by $V(t) = A \sin(2\pi Ft + \phi)$. Parameters of the central sinusoid of each beam are shaded in light gray.

As expected, the periodogram (Figure 3.5c) has a poor resolution. Furthermore, strong sidelobes from beam 1 interfere with beam 2, and combined sidelobes from beams 1 and 2 render the detection of the weak beam 3 uncertain without the a priori knowledge of its existence. The Blackman-Tukey spectrum (Figure 3.5d) is much smoother than the periodogram but the resolution further degrades. Beam 3 has also lost too much amplitude to be easily detected without a priori knowledge. With only one taper, beams 1 and 2 of the multitaper spectrum (Figure 3.5e) are only slightly better defined than in the Blackman-Tukey spectrum. This is not surprising since the multitaper method differs little from the Blackman-Tukey method when only one taper is applied. Beam 3 suffers from severe interference with sidelobes of the two stronger beams as in the case of the periodogram. With 2 tapers, the lack of resolution up to at least 3 Hz is such that the single-taper spectrum appears to be far better. As a whole, classical methods perform about equally poorly in this example. In particular, the multitaper method seems to have lost most of its strengths on this short data segment. The maximum entropy spectrum (FPE optimal order $M_0 = 50$) has however no problem (Figure 3.5f) in clearly resolving the three beams. The weak beam 3 is particularly well defined compared to classical methods, both in

width and in amplitude. If the model order is severely underestimated as when an order 20 is selected, a considerable smoothing is produced and beams 2 and 3 just disappear. In that case, the peak value corresponding to beam 1 shifts also toward a slightly higher frequency. This extreme case illustrates the importance of selecting an appropriate model order in MEM.

Marple (1982) has shown that for two complex sinusoids, the resolution of MEM is increasing when the signal-to-noise ratio increases. When the SNR is very low ($\text{SNR} \approx -30$ dB), the resolution of MEM tends toward the resolution of the Blackman-Tukey method, which is independent of the SNR. For a fairly high SNR, the resolution of MEM is however significantly higher than the resolution obtained with the Blackman-Tukey method.

The asymptotical properties of MEM have been investigated by Kromer (1970). His findings indicate that the MEM spectral estimate is asymptotically normal, and unbiased

$$\lim_{N \rightarrow \infty} E\{\hat{P}_{MEM}(f)\} = P_{X,X}(f). \quad (3.29)$$

The variance of the MEM estimate is given for large N (number of data points) and M (AR model order) by

$$\text{Var}\{\hat{P}_{MEM}(f)\} = \frac{2M}{N} P_{X,X}^2(f). \quad (3.30)$$

Asymptotically, the variance of the MEM estimate is approximately equivalent to that of the Blackman-Tukey estimate with a suitably chosen truncation length (Gersch and Sharpe, 1973).

Despite the existence of asymptotical properties, MEM suffers from the lack of a clearly defined variance in the important cases for which N and M are not large. Baggeroer (1976) has however proposed a method of determining confidence intervals for Yule-Walker MEM spectral estimates. When noise is added to the exact autocorrelation of a real sinusoid (Lacoss, 1971), the MEM estimates display much larger variations in the peak value than Blackman-Tukey estimates do. However, fluctuations of the area under the peak (the estimated power) are comparable to the fluctuations in the peak value obtained with the Blackman-Tukey method. With two sinusoids, MEM tends to suppress a lower peak relatively to a higher one. As a result, the ratio between the high and low peak values is greater than their power ratio (square of the power ratio). This effect is compensated by a broadening of the lower peak in such a way that the area under the peak (the integrated power) reflects the true power of the lower peak. It is not clear however whether this property of MEM remains valid or not in the case of continuous spectra. A method for obtaining

a power estimate without numerical integration has been presented by Johnson and Anderson (1978).

Despite their particular qualities, MEM and related methods have never gained popularity in seismology. This is probably related to the immense popularity of Fourier methods as a result of the existence of highly versatile FFT algorithms. The model order determination in MEM might also have been an obstacle to its more widespread use. MEM and autoregressive methods appear however to be highly relevant to several aspects of the analysis of seismograms. Various seismological applications are documented, ranging from characterization of seismograms and high-quality power spectrum estimation (Lacoss, 1971; Leonard and Kennett, 1999) to automatic onset picking methods of single-component and multi-component recordings as mentioned earlier (Morita and Hamaguchi, 1984; Takanami and Kitagawa, 1988; Kushnir et al., 1990; GSE/Japan/40, 1992; Leonard and Kennett, 1999; Sleeman and van Eck, 1999; Takanami and Kitagawa, 2003). Phase onset picking is probably the most important application of autoregressive methods in seismology today.

Power spectrum estimation in MannekenPix

Since MEM appears to have the important and rare qualities required for the design of an effective Wiener filter derived from short data segments, I selected this method to evaluate power spectra in MannekenPix. Among the various algorithms available, MannekenPix uses the classical Burg's algorithm implemented in the "Numerical Recipes in Fortran" (Press et al., 1992). The validity of this particular implementation is supported by an atmospheric angular momentum study (Penland et al., 1991), and results derived from the Numerical Recipes closely match also those obtained with the Signal Processing Toolbox of Matlab 6.0.

Routine "Memcof" evaluates the AR coefficients, and routine "Evlmem" evaluates the power spectral density at a given frequency. Once the AR coefficients have been determined for a given model order, successive calls to "Evlmem" allow to form a complete spectrum, one frequency per call.

Three methods of determining the optimal model order have been tested. In the first method, fixed orders were used, the optimal order being the model order maximizing a criterion that takes into account both the accuracy and the hit rate (the relative frequency of picked seismograms) of the picking for a whole data set. In the second method, Akaike's FPE criterion was used to select

an optimal model order for each seismogram. In the third method, the optimal model order was the order maximizing the signal-to-noise ratio after the application of several Wiener filters with different model orders for each seismogram.

A data set composed of more than 1,000 short-period seismograms from local earthquakes recorded by stations located in the vicinity of the Dead Sea basin, was carefully re-picked and re-weighted manually (see Chapter 4). The automatic picking of this reference set with MannekenPix revealed only a weak sensitivity of the global results to the model order for orders not greater than half the sample size. The third method, i.e. a model order derived from the optimal signal-to-noise ratio, appeared however to provide the best global results and was adopted as the standard method to select the model order in MannekenPix.

Application of the Wiener filter

Once the power spectrum of a segment of noise has been evaluated by MEM, it can be subtracted from the power spectrum of a data segment containing both noise and signal in the vicinity of the first-arrival time. The result is an approximation of the power spectrum of the P-wavetrain. The magnitude spectrum of the Wiener filter can then be computed by equation (3.15).

In order to avoid introducing an additional delay to the onsets by the application of a causal filter, MannekenPix uses a zero-phase finite impulse response (FIR) Wiener filter. FIR filters are perfectly suited for implementing the complex shapes of Wiener filters because their magnitude spectra can take nearly any shape, and their phase spectra can be made exactly linear or zero-phase (e.g. Oppenheim et al., 1999). In addition, FIR filters are always stable.

Since the Wiener filter only attenuates frequencies for which the amount of signal is low compared to the amount of noise, precursory signal sidelobes are usually not a concern with this noncausal filter, and the signal amplitude remains quasi-unaffected by the filtering process. Precursory noise sidelobes should well be produced but they contribute to the residual noise after filtering, and they do not have the special importance that signal sidelobes have within the framework of onset picking.

The Wiener filter is applied in MannekenPix by routine “convlv” (Press et al., 1992). This routine multiplies the FFT spectrum of the data by the FFT spectrum of the Wiener filter but it

requires the inputs to be defined in the time domain. Since the Wiener filter of MannekenPix is defined in the frequency domain, its impulse response is obtained by an inverse FFT.

The window techniques used to reduce the spectral leakage of power spectra can also be applied to the impulse response of the Wiener filter in order to achieve higher stopband attenuations. However, when narrow spectral features (a few 0.1 Hz wide) have to be canceled, the smoothing introduced by windowing reduces the attenuation instead of improving it. Hamming and Blackman windows have been tested but these windows did improve neither the accuracy nor the hit rate of the picking. Consequently, MannekenPix does not window the impulse response of the Wiener filter prior to its application.

Wiener filter and deconvolution applied to a synthetic example

In order to better understand how the various parts of the Wiener filter operate together in MannekenPix, I designed a synthetic example where real-world seismic noise is combined with a synthetic (P-wave) pseudo-wavetrain. Full details of how the example was build are provided in Appendix B. The idea however is not to produce a completely realistic simulation of a seismic record. I mainly want to show the typical distortion that a common recording system inflicts on whatever it records, and how a Wiener filter and deconvolution can partially restore the original signal. This process is quite general and it is as important to manual picking as it is important to automatic picking.

First, I generated a synthetic source function (Figure B.1a in Appendix B) according to Brune's model (Brune, 1970) with program Pitsa (Scherbaum and Johnson, 1992). The Brune's model generates an S-wave source function but the distinction is not important here since P- and S-source functions are not very different from each other. By convolving the source velocity function (Figure B.1b) with a series of spikes of decreasing amplitude (Figure B.1c), a noiseless pseudo-wavetrain of 3 seconds (Figure B.2a) length was produced.

Second, I extracted 10 seconds of noise from a seismogram of a local earthquake recorded in the Dead Sea region by a short-period seismograph of GII. In order to better approximate true ground noise, the recorded noise was deconvolved (Figure B.2b) by the response of a digital seismograph (Figure B.5) composed of a simulated digital Marks L4-C seismometer (Figure B.3) and a digital 5th-order Butterworth low-pass (anti-aliasing) filter (Figure B.4).

The pseudo-wavetrain was added to the background noise at time $t = 7.0$ sec (Figure B.2c). The resulting time series simulates 7 seconds of pre-arrival noise and the first 3 seconds of a very noisy P-wavetrain unaffected by the instrumental response. The filtering effect of the seismograph (seismometer and anti-aliasing filter) is the final step and results in the simulated recorded seismogram of Figure 3.6a. The first lobes of the pseudo-wavetrain are close to what first-arrivals look like on real-world seismograms. What comes later is however not realistic since it is not based on the propagation effects of waves through a layered medium. It is only made of slowly decaying echoes of the source function. This is why I call it a pseudo-wavetrain.

Figure 3.6a displays the simulated recorded seismogram, where the noise segment and the signal-and-noise segment used by the Wiener filter have been color-coded in red and purple respectively. These two segments are 2.0 seconds long each, and they are separated from each other by a double safety gap of 100 msec (Figure 3.6c) around the true arrival time $t = 7.0$ sec. The very high level of noise makes it impossible to accurately pick the onset time or even to confidently identify the polarity of first-arrival motion.

The power spectra of the unfiltered seismogram (Figure 3.7a) show that the power density of the noise is concentrated around 1.7 Hz while the power density of the estimated signal is concentrated around 10.0 Hz. The high level of the estimated signal between 0.1 and 1.2 Hz appears dubious, and results at least partially from an imperfect stationarity of the noise between the noise segment and the signal-and-noise segment. In addition, the very high level of noise adversely affects the accuracy of the estimation of the signal power spectrum.

The shape of the Wiener filter (Figure 3.7b) is however rather simple and consists mainly of three narrow rejection bands, among which the band centered at 1.7 Hz is the most important one since it largely contains most of the noise power. The plot also displays the magnitude spectrum of the un-windowed FIR filter truncated in the time domain to a length of 4 seconds, and the spectrum of the corresponding Hamming-windowed filter. The rejection band centered at 1.7 Hz is too narrow to benefit from windowing and the un-windowed filter provides a better attenuation of the noise in this narrow band. The windowed filter appears however to perform better in the two broader rejection bands, but the contribution of these bands to the total noise is minor compared to the low-frequency band. As a whole, the un-windowed filter appears thus to perform better.

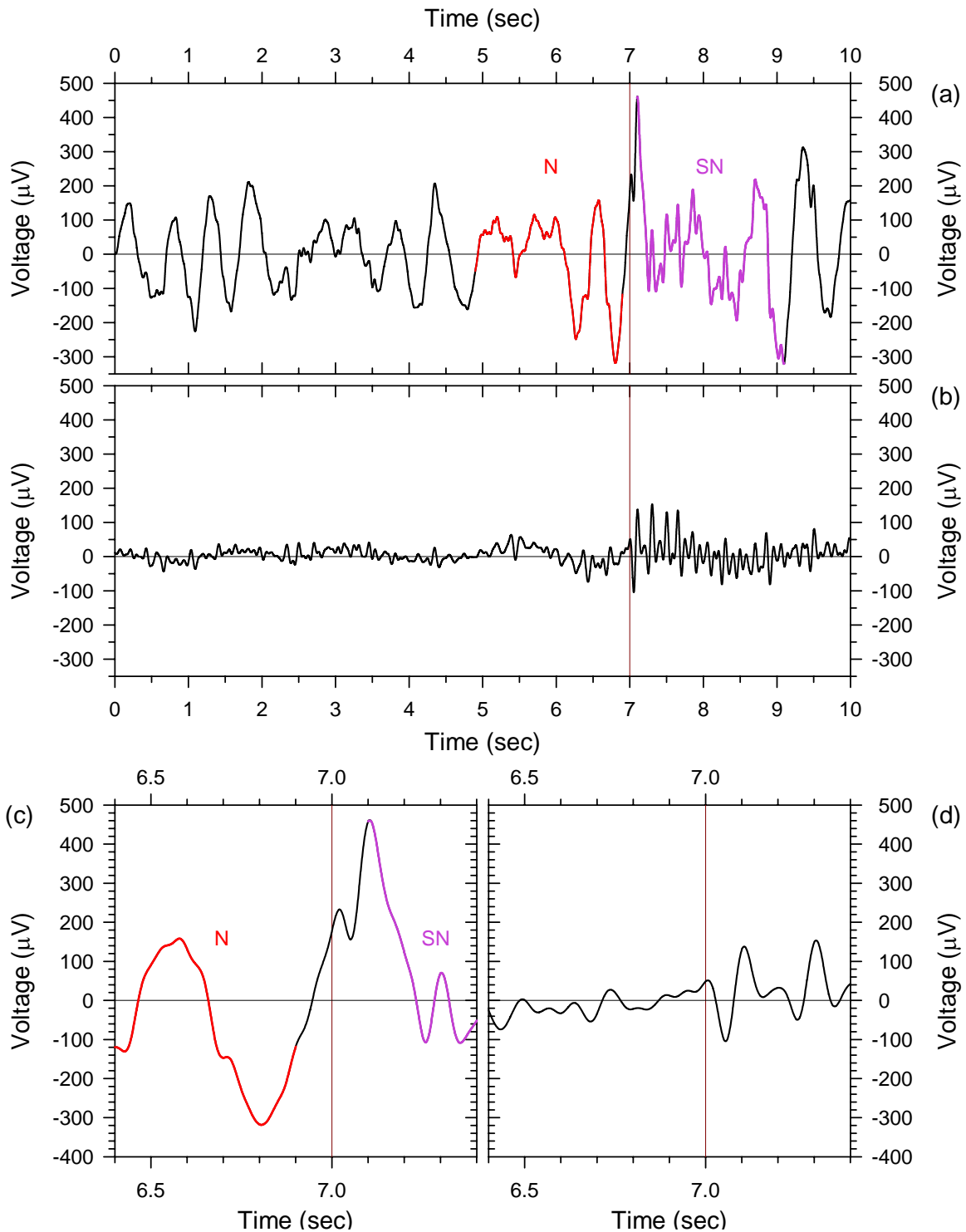


Figure 3.6 Simulated seismogram of a synthetic P-wave pseudo-wavetrain as it would be recorded by a short-period seismograph. (a) Raw seismogram (full). (b) Wiener-filtered seismogram (full). (c) Raw seismogram (detail). (d) Wiener-filtered seismogram (detail). The first arrival occurs at time $t = 7.0$ sec and is marked by the dark red vertical line. The 2-second noise (N) segment used in the evaluation of the Wiener filter is plotted in red (a,c). The 2-second signal-and-noise (SN) segment used in the evaluation of the Wiener filter is plotted in purple (a,c). The sampling period is 10 msec.

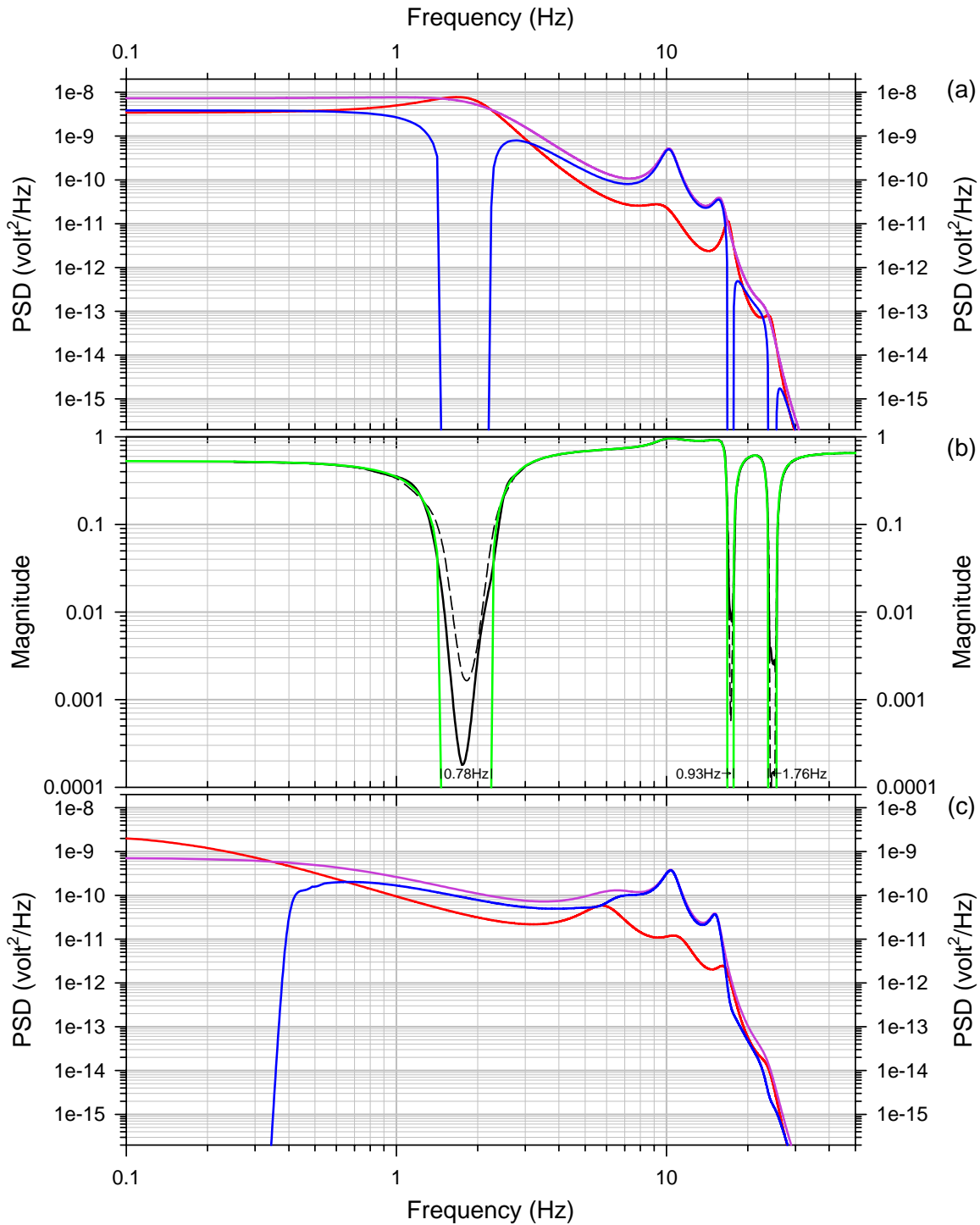


Figure 3.7 Power spectra and Wiener filter of the simulated seismograms plotted in Figure 3.6. (a) Power spectra before Wiener filtering (Noise in red, mixed Signal-and-Noise in purple, and estimated Signal in blue). (b) Wiener filter magnitude spectrum derived from power spectra plotted in (a). Ideal Wiener filter in green, actual filter (unwinded FIR of 4 sec.) as solid black, and windowed (Hamming) filter as dashed black. The 3 values at the bottom indicate the width (in Hz) of the narrow bands where a total cancellation of the noise is expected. (c) Power spectra of the Wiener-filtered seismogram. A MEM model order value of 9 has been selected, maximizing the signal-to-noise ratio of the Wiener-filtered seismogram.

The power spectra of the filtered seismogram (Figure 3.7c) reveal the presence of low-frequency (below 0.5 Hz) residual noise, easily observable on the seismogram (Figure 3.6b). This low-frequency noise has however very little incidence on the picking because its amplitude is now moderate compared to the amplitude of the signal, and the frequencies of the signal are much higher and quite distinct from those of the noise. One possible way to further reduce the residual noise would be to apply an additional Wiener filter to the already filtered seismogram. Such a procedure might prove beneficial to this particular example but it appears to degrade results when it is systematically used on large data sets.

The Wiener-filtered seismogram allows a far better identification of the first-arrival motion polarity (Figure 3.6b and d), and the uncertainty on the onset time is about ± 30 msec at the most (Figure 3.8a). A comparison between the noiseless recorded signal and the filtered seismogram (Figure 3.8a) demonstrates also that the Wiener filter does not delay the recorded signal. Due to residual noise, the example is however not conclusive about the absence of precursory sidelobes.

In order to partially correct the delay and distortion (Figure B.6) of the signal created by the seismograph, the Wiener-filtered seismogram was further deconvolved by the impulse response (Figure B.5) of the seismograph. The deconvolution was performed by spectral division in Matlab (see Appendix B), regularized by a small water-level value of 5×10^{-10} .

The result is plotted in Figure 3.8b along with the true signal, and it clearly demonstrates the potential importance of deconvolution for picking purposes. The most beneficial effect of deconvolution is perhaps its ability to re-focus the signal, smeared in this example by the awful phase properties of the Butterworth anti-aliasing filter. It is almost impossible to correctly identify the main features of the true signal from the Wiener-filtered seismogram (Figure 3.8a) alone. The filtered and deconvolved seismogram (Figure 3.8b) is also far from being perfect due to the residual noise but it accurately depicts several features of the true signal. For instance, the impulsive rise of the first arrival and echoes is rather well recovered and the sharp tips coincide exactly with those of the true signal. The first arrival and echoes displayed in Figure 3.8b are a bit broader than the true signal but they are clearly resolved, and they are also much closer to the true signal than on the Wiener-filtered seismogram that was not deconvolved (Figure 3.8a). The first-arrival onset time is also much easier to pick confidently. I would pick it at time $t = 6.990$ sec on the filtered and deconvolved seismogram, only one sample earlier than the true onset time.

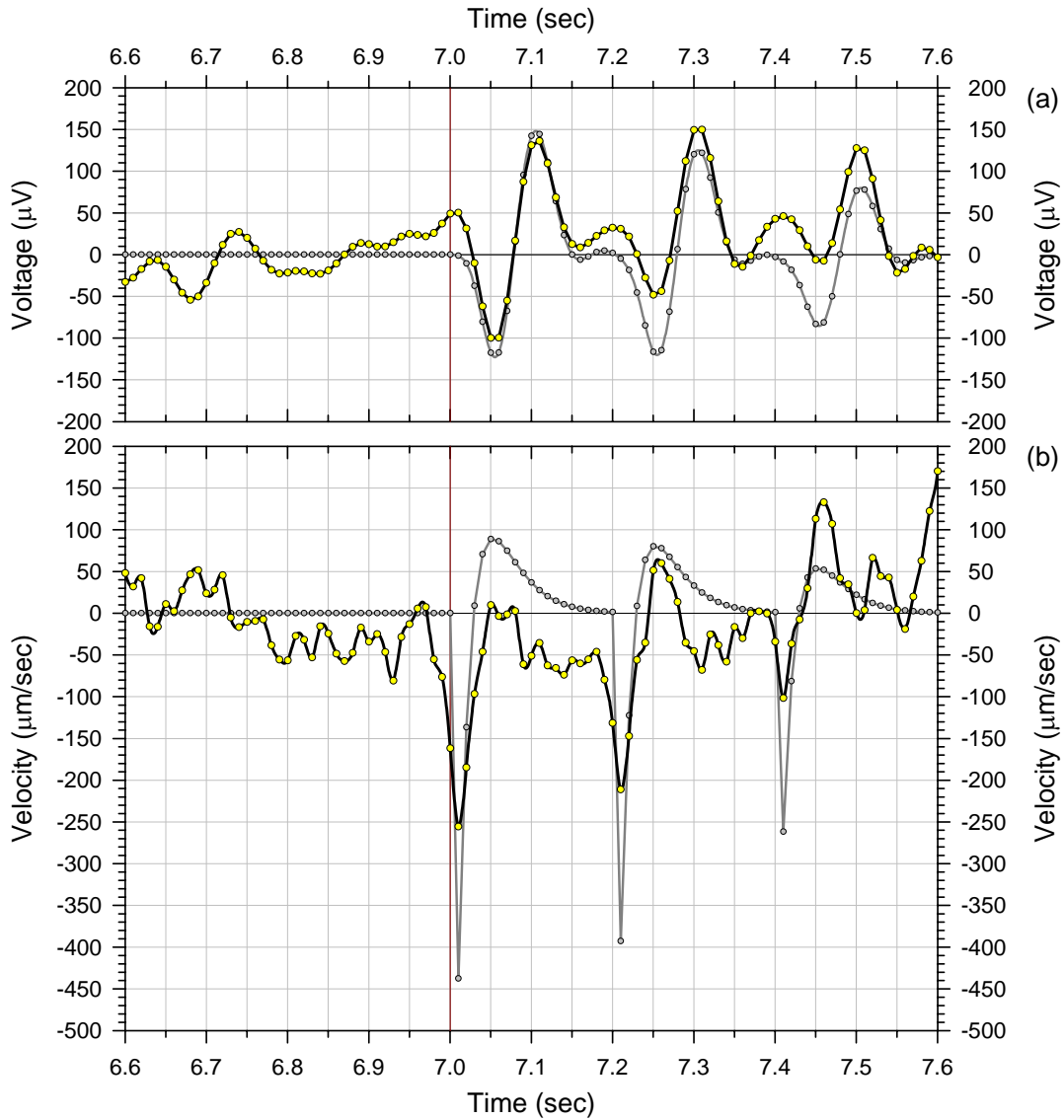


Figure 3.8 Deconvolution of the simulated seismogram by the seismograph response. (a) Wiener-filtered seismogram (black), and noiseless signal filtered (gray) by the seismograph response (Figure B.5a). (b) Wiener-filtered seismogram deconvolved (black) by the seismograph response, and noiseless true signal (gray). The sampling period is 10 msec.

3.3.2 AUTOMATIC PICKING

After the application of a Wiener filter in step 1 of MannekenPix, the automatic determination of the first P-wave onset time is done by the picking algorithm of Baer and Kradolfer (1987).

This algorithm defines first an approximate squared envelope function E_i^2 of the seismogram x_i , given by

$$E_i^2 = x_i^2 + \dot{x}_i^2 \frac{\sum_{j=1}^i x_j^2}{\sum_{j=1}^i \dot{x}_j^2} \quad (3.31)$$

where \dot{x} denotes the time derivative of x . The algorithm uses then a characteristic function CF_i , defined as

$$CF_i = \frac{E_i^4 - \overline{E_i^4}}{\sigma^2(E_i^4)} \quad (3.32)$$

in which $\overline{E_i^4}$ is the mean of E_i^4 from $j = 1$ to i and $\sigma^2(E_i^4)$ is the variance of E_i^4 . This characteristic function differs only from the statistical Z score of E_i^4 by the use of the variance σ^2 as denominator in (3.32) instead of the standard deviation σ used in the Z score.

A pick flag is set if CF_i increases above a given value *Threshold1*. The onset is accepted as a valid pick only if the value of the characteristic function stays above *Threshold1* for a certain amount of time *TUpEvent*. If the characteristic function drops below *Threshold1* after a shorter duration than *TUpEvent*, the pick is rejected and the pick flag is removed. The pick flag is however not removed if the characteristic function drops below *Threshold1* for an amount of time shorter than *TdownMax*. This happens frequently for events with low to moderate signal-to-noise ratios. The variance $\sigma^2(E_i^4)$ is continuously updated to account for variations of the noise level, except when CF_i exceeds the given value *Threshold2* usually chosen greater than *Threshold1*. The variance of E_i^4 is thus frozen when the validity of the onset is examined, shortly after the pick flag has been set.

In MannekenPix, the *Threshold1* value of the characteristic function that triggers the pick flag is determined in a fully adaptive and automatic way. For each seismogram, *Threshold1* is first set equal to the highest value of CF_i within the time window corresponding to the noise segment of the Wiener filter (e.g. Figure 3.6a). If the signal-to-noise ratio after the application of the Wiener filter is greater than 14 dB, the value of *Threshold1* is increased according to an internal table derived empirically. Regarding the amount of time *TUpEvent* required to validate an onset after a pick flag has been set, Baer and Kradolfer (1987) recommend to use at least one full cycle of the longest signal period expected. In the case of local events that occurred in the vicinity of the Dead Sea basin, *TupEvent* has an optimal value at 0.5 sec. A value of 1.0 sec could be a sensible choice when regional and local events are mixed together in one single picking set. The amount of time *TdownMax* during which CF_i can drop below *Threshold1* without clearing the pick flag has an optimal value in MannekenPix at half the signal period of highest power density as

measured by the Wiener filter routine. This value is quite consistent with Baer and Kradolfer (1987) who recommend half the mean of the two corner periods of their bandpass filter. $Threshold2$, the threshold value for freezing the variance update of E_i^4 , is optimal in MannekenPix at a value of $2 \times Threshold1$. This value is again in perfect agreement with the recommendations of Baer and Kradolfer. Another good value for the Dead Sea data is $1.1 \times Threshold1$, but a value of $2 \times Threshold1$ minimizes the variance of the automatic picking errors as determined from more than 1,000 reference picks.

Figure 3.9 displays the automatic picking result of the Wiener-filtered synthetic example of section 3.3.1. The plot was exported from program Visual Picker 1.2, in which most functions of MannekenPix are implemented with a graphical interface. MannekenPix was unable to find a valid onset for this example. The characteristic function triggers well the pick flag at time $t = 7.040$ sec but it does not stay long enough above the $Threshold1$ value to validate the pick.

3.3.3 DELAY CORRECTIONS

One known shortcoming (e.g. Sleeman and van Eck, 1999) of the Baer-Kradolfer algorithm is that the raw picking time provided is always somewhat late compared to what an analyst would determine as the onset time of a valid phase. For local earthquake data recorded by short-period instruments, this delay can be as small as one sample for the best seismograms. An example of such a small delay is provided in Figure 3.10b. For most seismograms, the delay is greater due to the interference of noise. Figure 3.11d displays a noisy seismogram where a longer delay is observed.

In order to reduce the delay of the Baer-Kradolfer onset times, all versions of MannekenPix include a primary delay correction. This correction derives from how the onset is determined by the Baer-Kradolfer algorithm. When a pick flag is set and validated, the characteristic function value is generally far above the background level observed in the segment of noise immediately preceding the onset. The idea for the correction is to move the raw automatic onset back in time as long as the characteristic function decreases significantly toward earlier samples. The delay correction stops when $(CF_i - CF_{i-1})$ is smaller than 0.01, or when this condition cannot be met after moving back the onset by 3 samples. This process is not perfect since the characteristic function may not decrease monotonically toward its background level, or because a clear background level may not even exist in the vicinity of the onset due to the

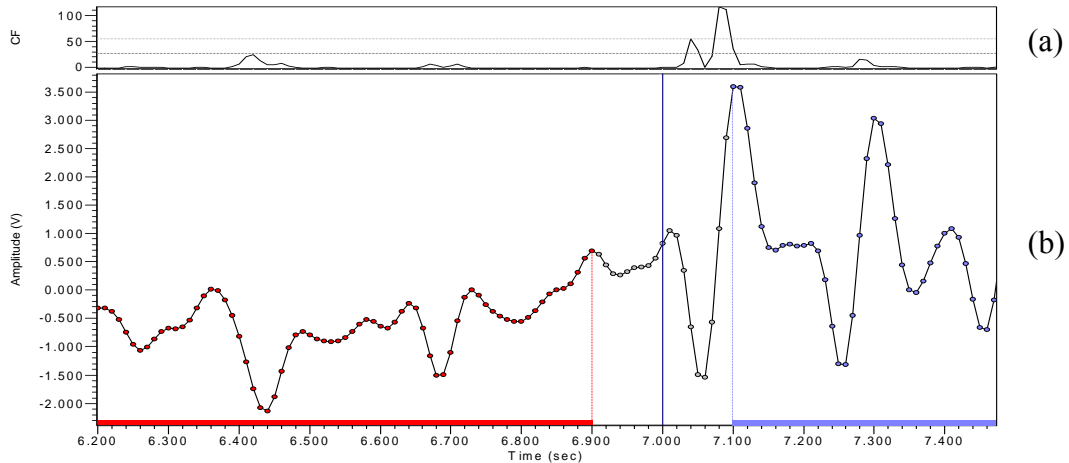


Figure 3.9 Automatic picking of the Wiener-filtered synthetic example (Figure 3.6b). (a) Characteristic function (CF). Lower dashed line is *Threshold1*, upper dashed line is *Threshold2*. (b) Wiener-filtered seismogram. Noise segment in red, Signal-and-Noise segment in blue. True arrival time in dark blue at time $t = 7.0$ sec. CF does not stay long enough above *Threshold1* to validate the pick flag triggered at time $t = 7.040$ sec.

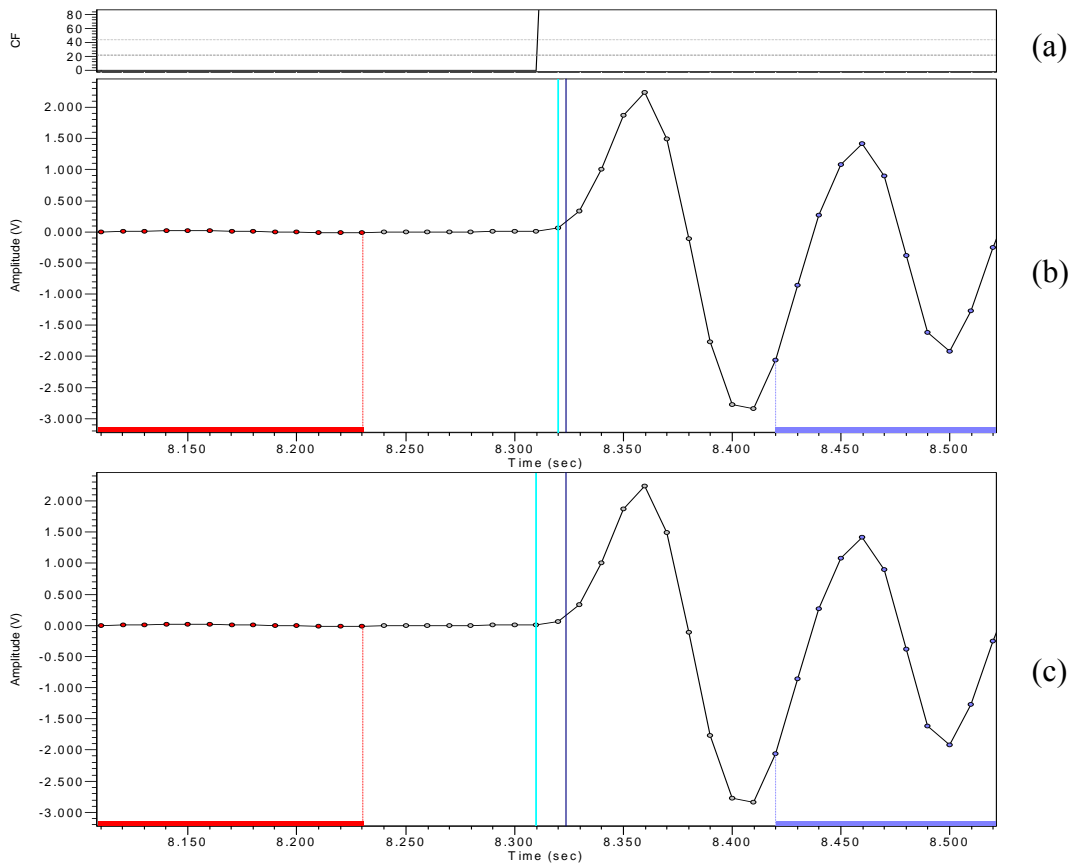


Figure 3.10 Automatic picking and delay correction of a high-quality seismogram from a local microearthquake that occurred at the Northern end of the Dead Sea basin (recorded by GII, Date 04 Dec 1986, Sta MOI, Vertical component short-period, O.T. 22:38:27.970, Dist 20.1 km, Depth 17 km, M_L 2.0, Sampling 10 msec). (a) Characteristic Function. Lower dashed line is *Threshold1*, upper dashed line is *Threshold2*. (b) Raw automatic pick (cyan), and routine pick (dark blue). (c) Delay corrected automatic pick (cyan), and routine pick (dark blue). (b,c) Noise segment in red, Signal-and-Noise segment in blue.

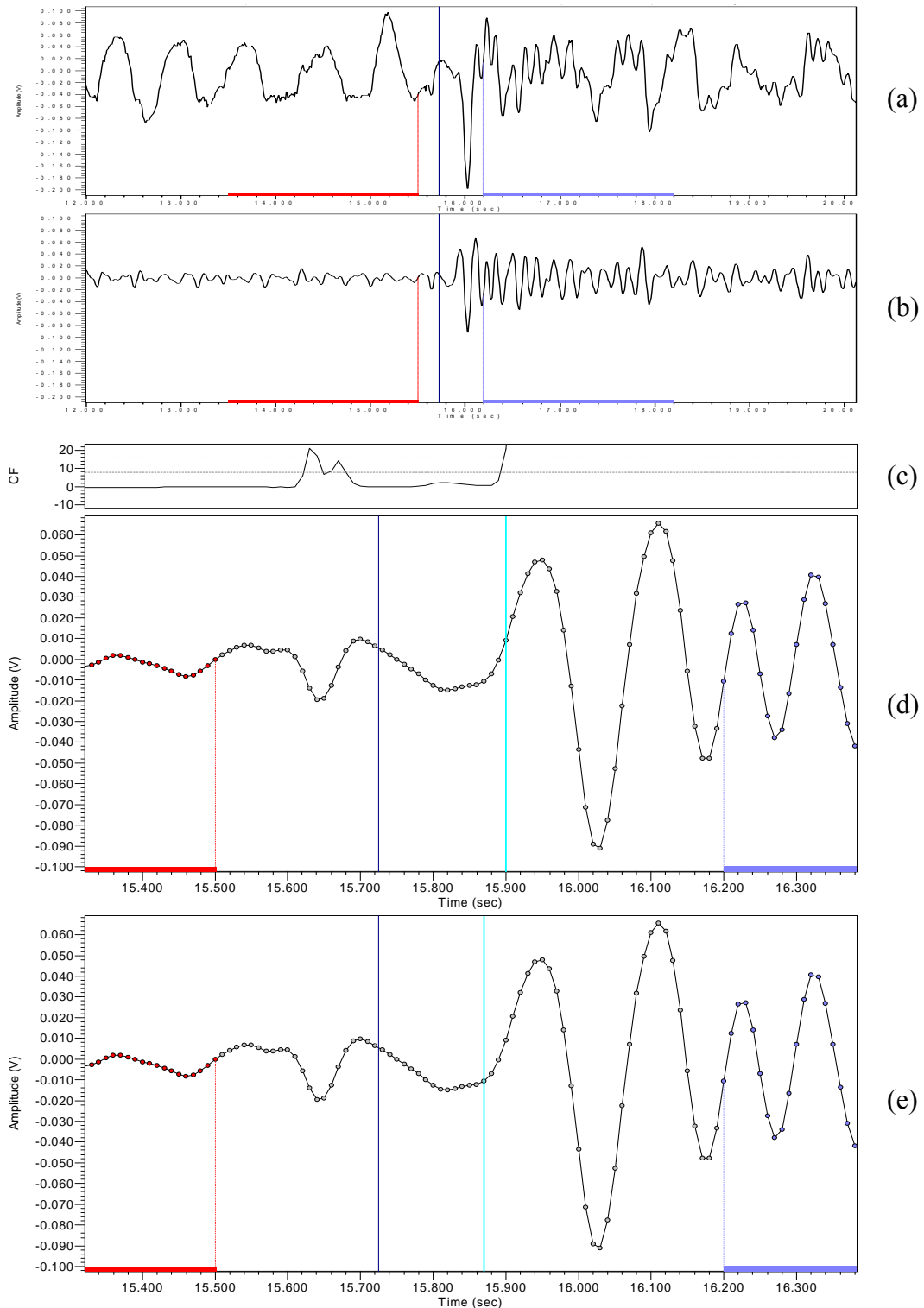


Figure 3.11 Automatic picking and delay correction of a noisy seismogram from a local microearthquake that occurred in the Northern part of the Dead Sea lake (recorded by GII, Date 11 Nov 1986, Sta DSD2, Vertical component short-period, O.T. 02:31:08.069, Dist 55.2 km, Depth 14 km, M_L 1.6, Sampling 10 msec). (a) Unfiltered seismogram. (b) Wiener-filtered seismogram. (c) Characteristic Function of (d,e). (d,e) Wiener-filtered seismogram (detail of b). (d) Raw automatic pick (cyan), and routine pick (dark blue). (e) Delay corrected automatic pick (cyan), and routine pick (dark blue). (a,b,d,e) Noise segment in red, Signal-and-Noise segment in blue.

disturbance by strong noise. Nevertheless, this simple correction usually provides good to very acceptable results for local earthquakes recorded by short-period instruments. Figure 3.10c and Figure 3.11e display the delay corrected onset times for the two examples presented before.

When regional and teleseismic data are picked with the original Baer-Kradolfer algorithm, raw picking delays can be much longer than a few samples. This behavior of the algorithm has also been observed by Sleeman and van Eck (1999). It results apparently from the fact that little or no frequency contrast (\dot{x}^2 terms in 3.31) between the P-wavetrain and the noise contributes then to the build-up of the characteristic function for these events. At lower frequencies typical of greater epicentral distances, the characteristic function is slow to reach the picking threshold level. The primary correction becomes then insufficient.

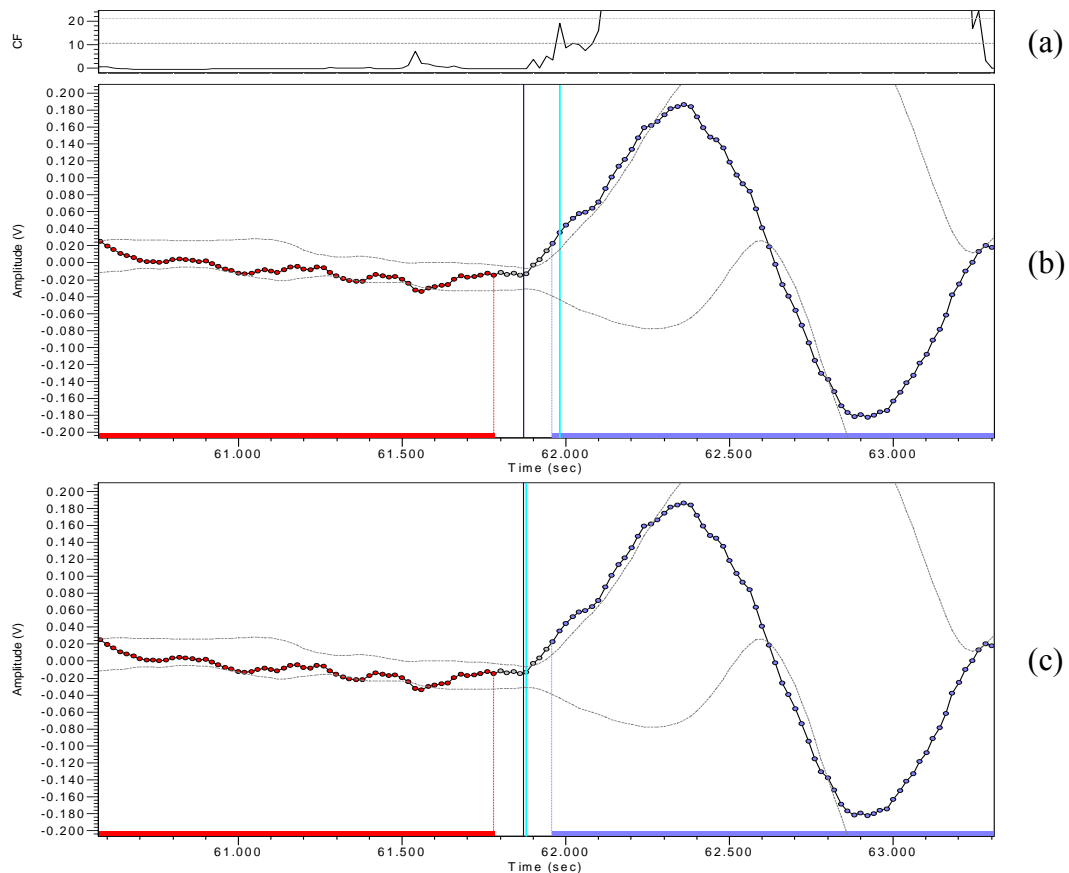


Figure 3.12 Automatic picking and deviation band delay correction of a seismogram from a regional earthquake that occurred in Turkey ~ 60 km N-W of Ankara (Recorded in Italy by INGV, Date 23 Aug 2000, Sta EB9 ~ 80 km W of Bologna, Vertical component short-period, O.T. 13:41:12.400, Dist 1,850 km, Depth 10 km, M 5.0, Sampling 20 msec). (a) Characteristic Function. Lower dashed line is *Threshold1*, upper dashed line is *Threshold2*. (b) Raw automatic pick (cyan), and routine pick (dark blue). (c) Deviation band delay corrected automatic pick (cyan), and routine pick (dark blue). (b,c) Noise segment in red, Signal-and-Noise segment in blue, and deviation band limits in gray.

In order to better correct the delay for regional and teleseismic arrivals, I introduced a secondary delay correction in version 1.7 of MannekenPix. This secondary correction is applied immediately after the primary correction. It does not appear to interfere adversely with the primary correction, so important for local earthquake data. An example of a first P-wave onset from a regional earthquake is presented in Figure 3.12b (from Istituto Nazionale di Geofisica e Vulcanologia, Roma). The earthquake occurred in Turkey near Ankara and was recorded near Bologna in Italy at an epicentral distance of about 1,850 km. The routine pick appears to be quite well located on the seismogram but the raw automatic onset time is late by about five samples. Since the characteristic function is not monotonically decreasing toward its background level when looking backward in time, the primary correction just applies an insufficient delay correction of one sample to the raw automatic time.

For non-impulsive events where the picking is primarily the result of an amplitude anomaly, a good delay correction is obtained by moving the pick earlier until a value of the seismogram becomes located inside a moving standard deviation band. A simple moving average $SMA_{i,P}$ of period P is given at the seismogram sample index i by

$$SMA_{i,P} = \frac{\sum_{j=0}^{P-1} x_{i-j}}{P}, \quad i \geq P. \quad (3.33)$$

The corresponding simple moving standard deviation $SMSTD_{i,P}$ is then

$$SMSTD_{i,P} = \left[\frac{\sum_{j=0}^{P-1} (x_{i-j} - SMA_{i,P})^2}{P} \right]^{1/2}, \quad i \geq P. \quad (3.34)$$

MannekenPix 1.7 uses a number of samples P in (3.33) and (3.34) corresponding to the period of highest noise spectral density as determined by the Wiener filter routine. The standard deviation band of MannekenPix 1.7 is defined as

$$STDB_{i,P} = SMA_{i,P} \pm 2 \times SMSTD_{i,P}, \quad i \geq P. \quad (3.35)$$

For a positive onset (increasing amplitude of first motion), the precise condition is that the delay correction stops at the first sample x_i below the higher value of $STDB_{i,P}$ in (3.35). This condition is usually met at an intermediate value between the two $STDB_{i,P}$ values. For a negative onset, the delay correction stops at the first sample x_i above the lower value of $STDB_{i,P}$. The

secondary correction is however skipped if the seismogram value x_i at the onset time corrected by the primary correction is not out the deviation band $STDB_{i,P}$ by more than $0.1 \times SMSTD_{i,P}$.

The raw automatic picking time of the regional earthquake of Figure 3.12b has been corrected by the primary and secondary delay corrections and the result is displayed in Figure 3.12c. The agreement between the manual pick and the final automatic pick is now excellent on this example.

3.3.4 WEIGHTING

In order to take into account the estimated uncertainty associated with picked phases, location and tomographic programs usually expect pick times to be accompanied by quality labels. In most applications, the highest quality is expressed by a small label value like 0 or 1 while the poorest quality is expressed by a higher value like 3, 4 or 5. These quality labels are then converted by the applications into observation weights that provide a variable importance to the arrival times according to their reported quality.

Uncertainties on arrival times are intervals of time. However, since the input of picking uncertainties into location and tomographic programs is not explicitly done as intervals of time, I am not sure that quality labels always reflect time uncertainties in practice. The use of quality labels instead of duration values appears to me as a choice favoring mispractices. It would have been far more in agreement with standard practices of the physical sciences to require the uncertainties to be expressed directly in seconds, or at the very least to require the quality labels to have the precise quantitative meaning they are supposed to have.

The weighting mechanism of the Baer-Kradolfer (1987) algorithm implemented in Pitsa (Scherbaum and Johnson, 1992) is based on the ratio between the P-wave amplitude within a certain time window and the largest amplitude of the seismogram before the pick flag is set for the first time. This kind of signal-to-noise ratio is certainly important for the weighting of arrival times, but I think that any variable by itself cannot carry all the information necessary to properly discriminate among time uncertainty groups. The weighting mechanism included in the Allen (1978, 1982) picking system is a great quality of this older system, and it uses four or five variables (depending on the implementation) to provide a pick weight. With a higher number of variables, it is potentially more powerful than the weighting mechanism included in Pitsa. One limitation of the Allen weighting mechanism is however that absolute thresholds (450 mV,

200mV, ...) for the amplitude of the P-wave intervene in the classification process. This might be a good choice for dense networks like the ones located in California, but it is certainly introducing a pessimistic bias for sparser networks where less redundancy is available. A common shortcoming of these two weighting mechanisms is that very little is known about the relation between their quality labels and time uncertainties. The core of the problem is not the internal consistency of these weighting mechanisms, it is how well the quality labels they provide correlate with time uncertainties on the readings.

MannekenPix includes a weighting mechanism where picking uncertainties are expressed as standard labels 1 to N , with N not greater than 5. A precise uncertainty duration defined by the user is however associated with each of these labels. This weighting mechanism is based on the discriminating power provided by the simultaneous use of several predicting variables as in the Allen picking system. Adapting the weighting mechanism of the Allen system to particular networks and data requires however a lot of trial-and-error at the very least. In MannekenPix, a calibration is made according to a statistical technique known as *discriminant analysis* on reference picks and associated weights supplied by the user. This calibration cannot be performed internally by MannekenPix at the moment, but it can be easily performed from the automatic picking output of the reference data set with a package like SAS, SPSS, or equivalent.

Discriminant analysis

Discriminant analysis is a statistical technique whose general purpose is to identify quantitative relationships between two or more criterion groups and a set of discriminating variables also called predictors. It requires the criterion groups to be mutually exclusive, and it assumes the absence of collinearity among discriminating variables, the equality of population covariance matrices and multivariate normality for each group. The mathematical objective is to weight and linearly combine the predictors in a way that maximizes the differences between groups while minimizing differences within groups (Fischer, 1936, 1938). When more than two groups are involved, the technique is commonly referred to as multiple discriminant analysis (MDA). As a *descriptive* technique, discriminant analysis serves to explain how various groups differ, what the differences between and among groups are on a specific set of discriminating variables, and which of these variables best account for the differences. As a *predictive* technique, it is used to predict the unknown group membership of cases based on the actual value of the discriminating variables.

The problem at hand in the weighting mechanism of MannekenPix is to predict the uncertainty group of automatic pick times from quantitative variables evaluated by the program. In order to be able to perform these predictions, it is however necessary to derive classification rules from an initial descriptive discriminant analysis on data for which the group membership is known. For a given data set, this discriminant analysis provides valuable guidelines to the user regarding both the definition of appropriate uncertainty groups and the selection of the most relevant predictors. The insight provided by the canonical discriminant functions also assists in the development of the weighting mechanism of MannekenPix by helping with the selection of a pool of standard predictors. The group membership predictions on unseen data are not done in a statistical package. MannekenPix performs this prediction internally by computing Fischer's linear discriminant functions from the coefficients determined during the initial descriptive discriminant analysis.

The discriminating variables of MannekenPix

The lack of discriminating power resulting from the use of only one predictor can be easily understood. For instance, any signal-to-noise ratio (SNR) derived from a segment of signal and a segment of noise implies already that a limiting choice has been made. Basically, the length of the segments can be chosen to be long or short compared to the dominant period of the signal. If long segments are selected, chances are that a general characterization of the overall quality of the seismograms will be gained. This overall quality might then discriminate rather well between picks located far away from true onsets and those located closer from true onsets. What is missing however in the prediction process is a localized SNR value derived from very short segments. Used in isolation, this localized SNR might reflect the accuracy of the picked features, but little will be known about the possibility of gross errors. When the information provided by both predictors is combined, some discriminating power is gained because a decision matrix can then be derived with information about both the possibility of gross errors and a very localized quality factor. Similarly, the discriminating power can be further increased by adding appropriate variables in the prediction process. Naturally, defining the most appropriate predictors is not a straightforward task.

The benefits of using discriminant analysis for this purpose are that trial-and-error is reduced to a minimum, no decision matrix needs to be coded explicitly, and the best predictors for specific data sets can be easily selected from the pool of standard discriminating variables already gathered. Discriminant analysis is thus useful both to define new predictors, and to select the

available ones most appropriate to tackle specific data sets. The pool of standard predictors available so far includes 9 variables evaluated internally by MannekenPix. On specific data sets, the optimal number of predictors usually ranges from 6 to 8 among the 9 available.

Predictor 1 (*WfStoN*) is a signal-to-noise ratio derived from the signal and noise power spectra determined by the Wiener filter routines. Its value is evaluated at the final automatic onset time and it is given in decibels by

$$P_1 = 10 \log_{10} \frac{\sum_{f=0}^{F_{Ny}} P_{S,S}(f)}{\sum_{f=0}^{F_{Ny}} P_{N,N}(f)} \quad (3.36)$$

where $P_{S,S}(f)$ and $P_{N,N}(f)$ are the power spectral densities of the signal and the noise respectively, and F_{Ny} is the Nyquist Frequency. Although the Wiener filter uses short data segments (about 2 sec for local data), the length of these segments is long compared to one apparent cycle of the P-wavetrain. Predictor 1 is an overall quality estimate around the final automatic onset time.

Predictor 2 (*GdStoN*) is a signal-to-noise ratio determined by the delay correction routines. Its value is evaluated at the final automatic onset time and it is given in decibels by

$$P_2 = 20 \log_{10} \frac{\sum_{f=0}^{F_{Ny}} A_S(f)}{\sum_{f=0}^{F_{Ny}} A_N(f)} \quad (3.37)$$

where $A_S(f)$ and $A_N(f)$ are the magnitude spectra of the signal and the noise respectively, and F_{Ny} is the Nyquist frequency. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the P-wavetrain. Predictor 2 is a localized measure of quality at the final automatic onset time.

Predictor 3 (*GdAmpR*) is a signal-to-noise ratio determined in the time domain by the delay correction routines. Its value is evaluated at the final automatic onset time and it is given in decibels by

$$P_3 = 20 \log_{10} \frac{Amp_S}{Amp_N} \quad (3.38)$$

where Amp_S and Amp_N are the peak-to-peak maximum amplitude of the signal and noise respectively. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the P-wavetrain. Predictor 3 is also a localized measure of quality at the final automatic onset time.

Predictor 4 ($GdSigFR$) is the signal-to-noise ratio evaluated at the dominant frequency of the signal, as determined by the delay correction routines. Its value is evaluated at the final automatic onset time and it is given in decibels by

$$P_4 = 20 \log_{10} \frac{A_S(F_S)}{A_N(F_S)} \quad (3.39)$$

where $A_S(F_S)$ is the value of the magnitude spectrum of the signal at the dominant frequency of the signal $f = F_S$, and $A_N(F_S)$ is the value of the magnitude spectrum of the noise at frequency $f = F_S$. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the P-wavetrain. Predictor 4 is also a localized measure of quality at the final automatic onset time.

Predictor 5 ($GdDelF$) is the difference between the dominant frequency of the signal and the dominant frequency of the noise, as determined by the delay correction routines. Its value is evaluated at the final automatic onset time and it is given in hertzs by

$$P_5 = F_S - F_N \quad (3.40)$$

where $f = F_S$ is the dominant frequency of the signal and $f = F_N$ is the dominant frequency of the noise. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the P-wavetrain. Predictor 5 is a localized measure of the frequency contrast between the signal and the noise.

Predictor 6 ($ThrCFRat$) is the natural logarithm of the ratio between the first maximum of the characteristic function after picking to the $Threshold1$ value, as determined by the automatic picking routines. It is given by

$$P_6 = \ln \frac{CF_{Max1}}{Threshold1} \quad (3.41)$$

Predictor 6 is a localized measure of quality derived from the characteristic function.

Predictor 7 (*PcAboThr*) is the percentage of characteristic function samples above the *Threshold1* value before the picked onset, as determined by the automatic picking routines. The data segment has the length of the Wiener filter noise segment and stops one sample before the picked onset. A Predictor 7 value different from zero is often associated with a possible earlier pick time, or with a high level of pre-arrival noise.

Predictor 8 (*PcBelThr*) is the percentage of characteristic function samples below the *Threshold1* value after the picked onset, as determined by the automatic picking routines. The signal data segment length is 1/10 the length of the Wiener filter mixed signal-and-noise segment and starts at the picked onset. A Predictor 8 value different from zero is often associated with low to moderate signal-to-noise ratios.

Predictor 9 (*CFNoiDev*) is a measure of deviation of the characteristic function before the picked onset, as determined by the automatic picking routines. It is given by

$$P_9 = \ln \frac{\sum_{i=1}^N |CF_i - CF_{median}|}{N |Threshold1 - CF_{median}|} \quad (3.42)$$

where the CF_i are the N samples of the characteristic function in a noise segment of length equal to the length of the Wiener filter noise segment, and CF_{median} is the median value of CF_i over the N samples.

Chapter 4

High-quality automatic picking of local earthquake data from the Dead Sea region

The automatic picking system MannekenPix was initially developed to collect a high-quality set of picking data from local earthquakes that occurred in the vicinity of the Dead Sea basin. This data set would then be used in another work to perform a high-resolution travel-time tomographic study of the crustal structure of the region.

Figure 4.1 displays the set of 531 well-locatable (≥ 8 P Picks and Gaps $< 180^\circ$) earthquakes recorded by GII during the period 1984-2001 for which digital seismograms could be retrieved. The yearly distribution of seismograms is shown in Figure 4.2. The data set is complete for the period 1984-1996 but a final check is required for the period 1997-2001, and it does not include any earthquake after September 2001. Eighty short-period stations (L4-C and some S-13) operated by GII and 10 short-period stations operated by JSO carry first-arrival P-readings. The coordinates of the 90 stations are provided in Appendix C. Most stations are vertical single-component stations, and only the vertical component has been used for the few three-component stations available. All the data were recorded by GII, mainly at a sampling rate of 50 Hz, but some earthquakes were recorded at 100 Hz in 1986, 1987, and 1993.

All seismograms of a specific event are stored by GII in a single binary file. I wrote a program to extract the seismograms from GII files and convert them to binary SAC format as required by MannekenPix. The routine picking information has been written into the SAC headers from related phase files. When no routine pick is available for a particular station, MannekenPix uses the predicted arrival computed by the conversion program as the a priori time around which to search for a P pick. The conversion program uses ray-tracing with model Israel (see 2.3) to compute these predicted values.

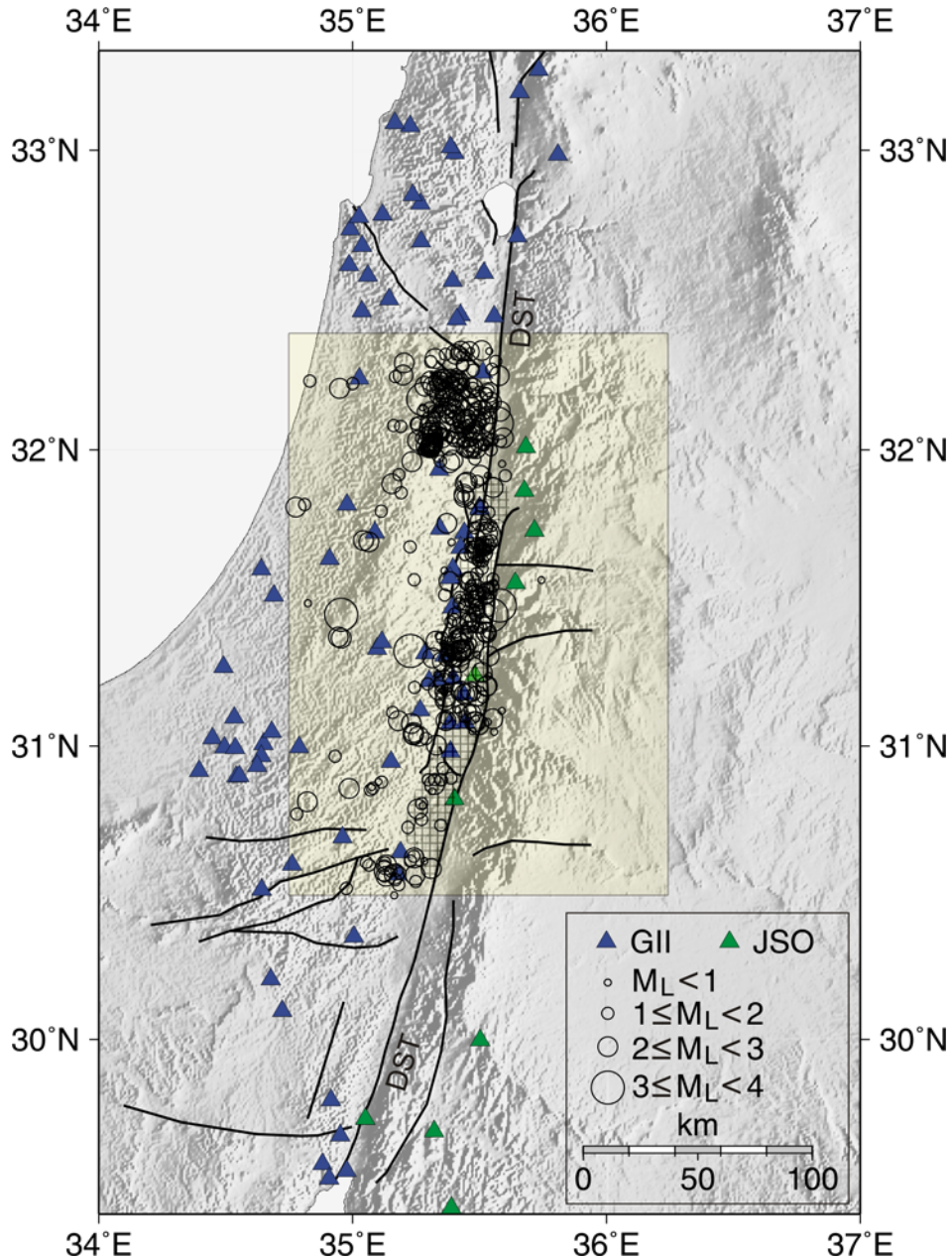


Figure 4.1 Five hundred thirty-one well-locatable earthquakes (1984-2001) in the vicinity of the Dead Sea basin recorded by short-period stations (triangles) of GII (Israel) and JSO (Jordan). The square grid fill defines the Dead Sea basin, and the earthquake selection area is in yellow. DST: Dead Sea Transform. Topographic shaded relief derived from Globe 1-km elevation data.

4.1 METHODOLOGY

The complete data set includes 15,250 seismograms, from which two subsets of particular importance have been extracted (Figure 4.2). The first subset is the *calibration* set. It is used to

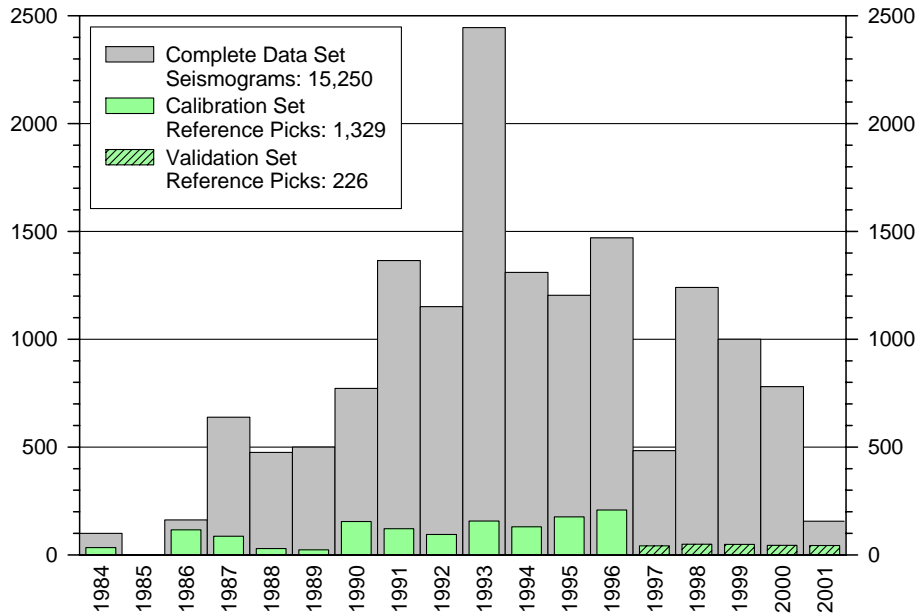


Figure 4.2 Yearly distribution of the number of seismograms and reference P picks.

determine the appropriate picking parameters, and to calibrate the weighting engine of MannekenPix from user-supplied reference picks and related uncertainties. The calibration set includes 1,329 reference picks and weights from myself, and it extends over the years 1984 to 1996. In order to evaluate the stability of the uncertainty predictions on totally unseen data, a small out-of-sample or *validation* set of 226 reference picks has also been created from seismograms spanning the 1997-2001 period.

The calibration and validation reference picks were performed with program Visual Picker 1.2 in which most functions of MannekenPix are implemented with a graphical interface. Since the best automatic picking results are expected from Wiener-filtered seismograms, the reference picking was also done on Wiener-filtered seismograms. This does not imply that a bias in favor of the automatic picking has been introduced, it means that a common de-noising technique has been used in both cases. This is extremely important because a precise calibration of the weighting engine is only possible if the human analyst and MannekenPix see sufficiently similar seismograms. In order to ease the visual quantification of the uncertainties, Visual Picker plots the uncertainty intervals on the seismograms during the picking process (Figure 4.3). All the reference picking was done with picking weights ranging from 1 to 5 (Table 4.1) but no weight 5 picks were included in the reference sets. Poor discriminating results from MannekenPix between picking weight 1 and weight 2 required the merging of these two groups into a single

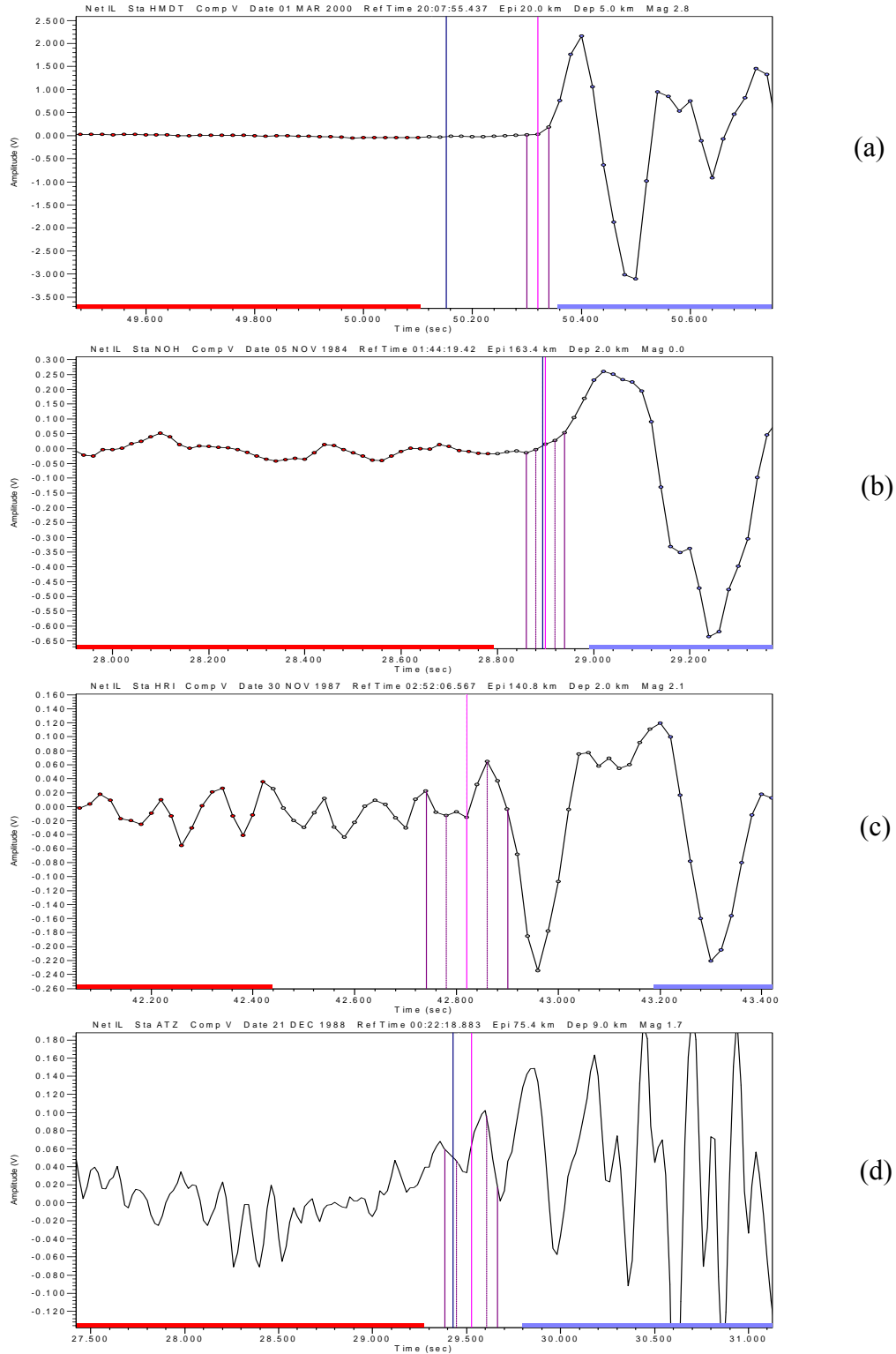


Figure 4.3 Examples of Picking Weights 1-4. (a) Picking Weight 1. (b) Picking Weight 2. (c) Picking Weight 3. (d) Picking Weight 4. Reference picks in magenta, routine picks in dark blue. The external solid lines (purple) on both sides of the reference picks delineate the actual uncertainty limits while the finely dotted inner lines (purple) delineate the limits of the previous (lower uncertainty) weight group.

weight 1 group. The weights used by MannekenPix for the Dead Sea area are the target weights reported in Table 4.1.

Table 4.1 Uncertainty groups

Uncertainty ε (ms)	Picking Weight	Target Weight ¹
$0 \leq \varepsilon \leq 20$	1	1
$20 < \varepsilon \leq 40$	2	2
$40 < \varepsilon \leq 80$	3	3
$80 < \varepsilon \leq 140$	4	4
$ \varepsilon > 140$	5	5

¹ Only the target weight is used for calibration and prediction.

The automatic filtering and picking of the complete data set with MannekenPix 1.7.6 took on the average 0.4 sec per seismogram (2.5 seismograms / sec, or 9,000 seismograms / hour) with a 1.4 GHz Pentium 4 processor, and Linux as operating system.

4.2 DATA SET CHARACTERIZATION

As mentioned in 3.1, a description of the main properties of automatically picked data sets is essential to potential users since by analogy with known cases they can then better decide if the technique might be useful to them. This is very easy to do with MannekenPix since the program provides various signal-to-noise ratios along with dominant frequencies for the signal and for the noise.

Figure 4.4 displays the distance distribution for the complete data set and for the calibration set. The range of epicentral distances for the complete data set (Figure 4.4a) is about 0-340 km. The seismicity occurs throughout the crust (0-32 km) but the mantle is apparently aseismic. The routine picking hit rate is 45.2 % of the total number of seismograms and routine picking extends about as far as the seismogram distance range. The number of seismograms culminates between 80 and 90 km while the routine picking rate culminates at closer range, between 40 and 50 km. The automatic picking hit rate is 46.5 % of the complete data set and it slightly exceeds (+1.3 %) the routine picking rate. The distance distribution differs however more clearly between the routine picking and the automatic picking. There are more routine picks between 0 and 90 km

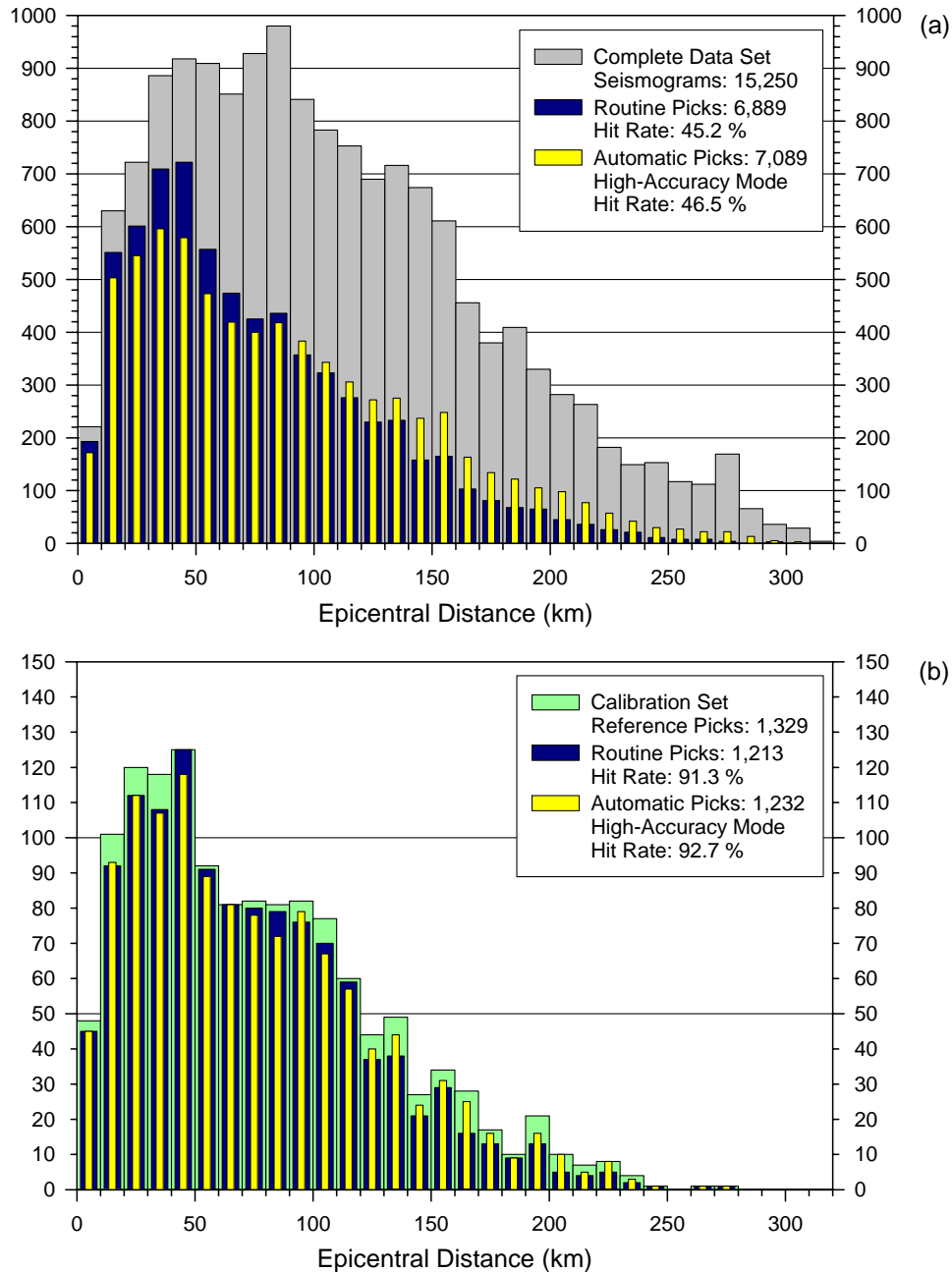


Figure 4.4 Distance distribution of the number of seismograms and P picks. (a) Complete data set. (b) Calibration set extracted from the complete data set.

but there are more automatic picks between 90 and 340 km. Figure 4.4b shows that the calibration set is rather similar to the routinely picked seismograms of the complete set in terms of distance range. The range is however a bit shorter as it does not exceed 270 km. On the other hand, the calibration set is smoother than the complete data set in terms of distance incidence on the hit rate. The average routine picking hit rate is 91.3 % of the reference picks while the

average automatic picking rate is 92.7 %. The automatic picking slightly exceeds (+1.4 %) here also the routine hit rate.

An overall Signal-To-Noise ratio measured on the unfiltered seismograms is displayed for the complete data set in Figure 4.5a. This is a very simple RMS signal-to-noise ratio evaluated in the time domain around the routine pick times. It represents mainly the SNR of unfiltered seismograms that can be picked. Large safety gaps between the mixed signal-and-noise estimation window and the noise estimation window prevent against contamination from routine picking errors. The mean SNR is 11.0 dB with a standard deviation of 12.5 dB. Some very low SNR (-20 to -40 dB) are not erroneously reported. The frequency contrast between high-frequency signal and low-frequency noise of high-amplitude is sometimes sufficient to allow some manual picking at these very low SNR.

Figure 4.5b displays the dominant frequencies of noise and signal evaluated around the routine pick times. The dominant frequencies of the noise were determined on unfiltered seismograms while the dominant frequencies of the signal were determined on Wiener-filtered seismograms. Large safety gaps were also used here. The dominant noise is below 2 Hz on most seismograms and the number of seismograms decreases regularly above 2 Hz until a background level reached at about 5 Hz. The background level of noise is low but non-neglectable until 25 Hz. Since dominant frequencies contain the highest spectral power for a given data segment, this means also that an adaptive Wiener filter is preferable for this data set and not only a high-pass Wiener filter as suggested by Douglas (1997). Dominant frequencies of the signal culminate between 5.5 and 6.0 Hz and decrease in a triangular way on both sides of the modal value. The signal extends also toward higher frequencies, but in a less pronounced way than the noise does.

The impact of the Wiener filter on the SNR is displayed in Figure 4.6. The unfiltered calibration set has a mean SNR of 16.3 dB and a standard deviation of 11.2 dB. This mean value is about 5 dB higher than the mean value of the complete data set. The fact that the SNR was estimated for the complete set at the routine time while the impact of the Wiener filter is estimated at the reference time is not an issue. Both means are within 0.5 Hz and their standard deviations differ by less than 10 %. The global impact of the filter is to increase the mean SNR by 4.3 dB and by reducing the standard deviation by about 25 %. Naturally, the effect of the Wiener filter is on the average more pronounced for low SNR than for high SNR, and the distribution of filtered seismograms is deeply skewed toward higher SNR. The three-parameter Gompertz regression displayed in Figure 4.6b provides an estimate of the SNR increase as a function of the unfiltered SNR.

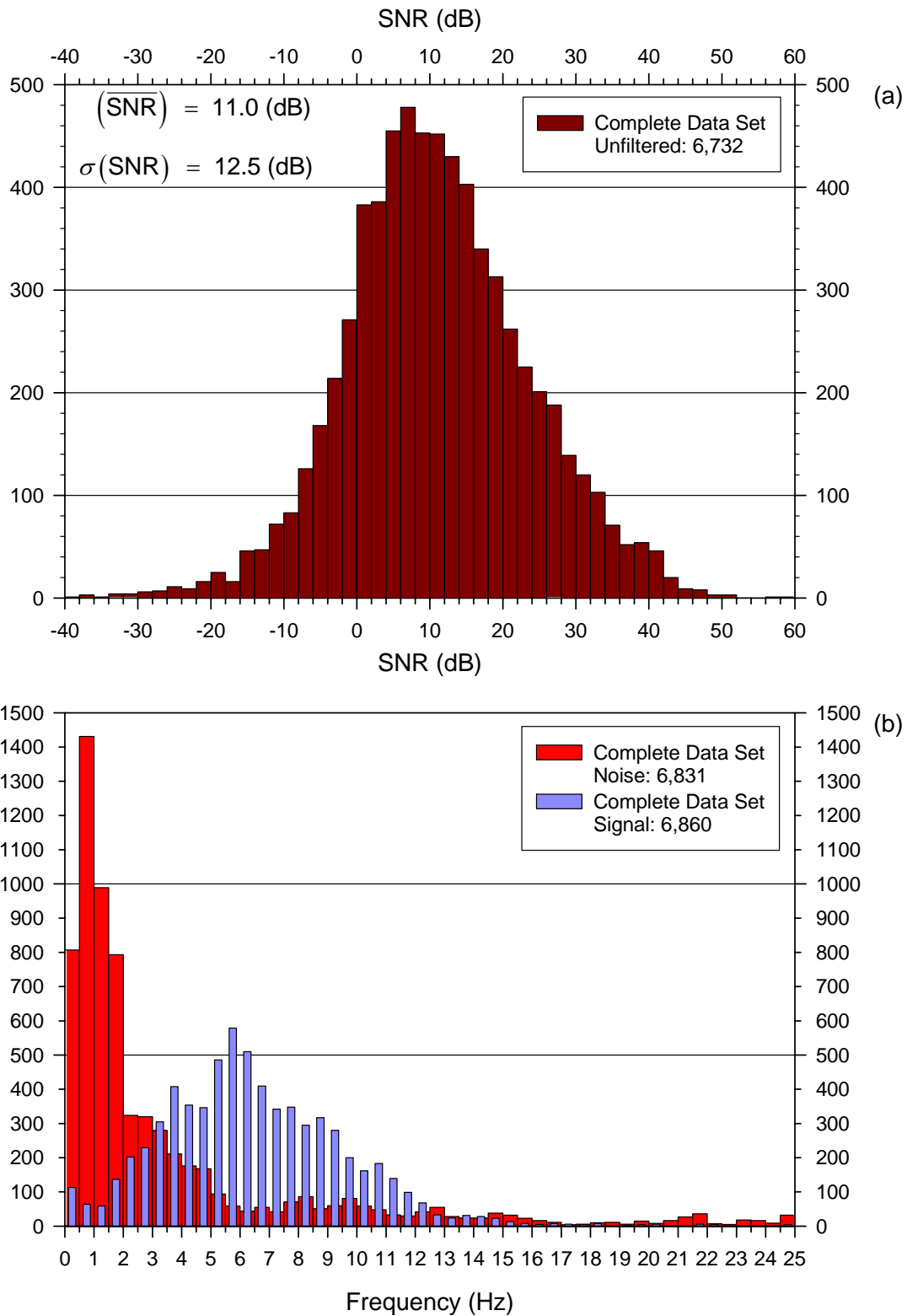


Figure 4.5 Signal-To-Noise ratio and dominant frequency distributions for the complete data set. (a) SNR distribution at the routine P picks from unfiltered seismograms. (b) Dominant frequency of Noise (from unfiltered seismograms) and dominant frequency of Signal (from WF seismograms) distributions at the routine P pick times.

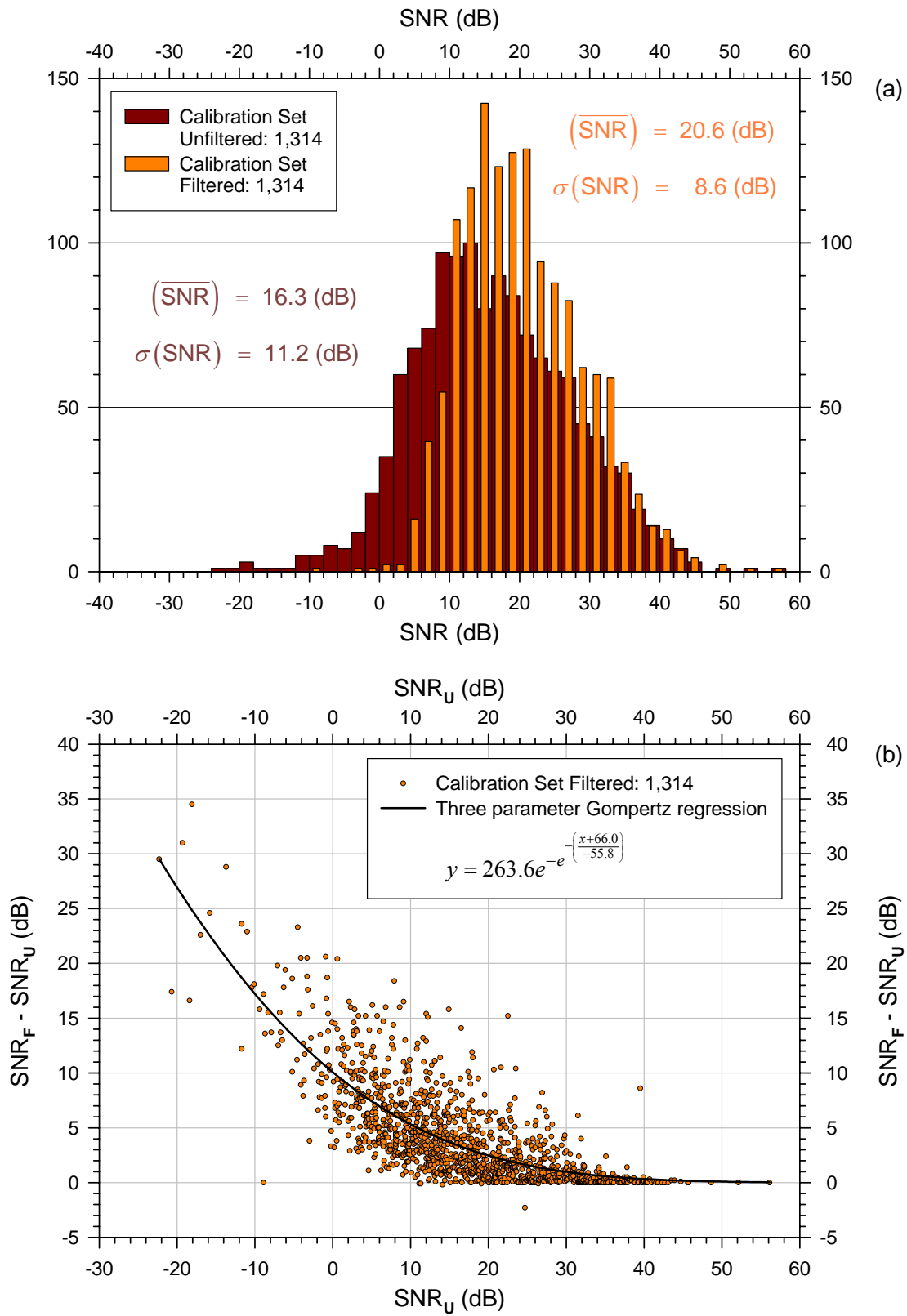


Figure 4.6 Impact of the Wiener filter on the Signal-To-Noise ratio in the calibration set. (a) Unfiltered and filtered SNR estimated at the reference pick times. (b) SNR increase by filtering as a function of the SNR of the unfiltered seismograms. SNR_U and SNR_F are the SNR of the unfiltered and filtered seismograms respectively.

4.3 PICKING RATE AND ACCURACY

Table 4.2 summarizes the picking rate on the three data sets. To speed up the reference picking of the validation set, only seismograms were selected for which both a routine pick and an automatic pick were known to exist from the automatic picking of the complete set. This explains why a picking rate of 100 % is found on this set for the routine and automatic filtered pickings. The validation set is not used to test the picking rate or the picking accuracy, it is only used to test the predictive power of the calibrated weighting engine on unseen data. Since Visual Picker does not display the automatic picking results obtained with MannekenPix, no picking bias in favor of the automatic picking times or weights is introduced by the selection process described here.

Table 4.2 Picking hits and hit rate

Set	Seismograms	Reference picks	Routine picks (Hit Rate)	Auto picks ¹ WF ² (Hit Rate)	Auto picks ¹ UF ³ (Hit Rate)
Calibration	1,329	1,329	1,213 (91.3 %)	1,232 (92.7 %)	1,098 (82.6%)
Validation	226	226	226 (100.0 %)	226 (100.0 %)	183 (81.0 %)
Complete	15,250	1,555	6,889 (45.2 %)	7,089 (46.5 %)	5,216 (34.2 %)

¹ High-Accuracy Mode.

² Wiener-Filtered.

³ Unfiltered.

As said earlier, the automatic picking hit rate on Wiener-filtered seismograms is slightly higher than the routine hit rate. On the complete data set, there is a drop of about 25 % in the number of automatically picked seismograms when the Wiener filter is not applied (Table 4.2). This outlines the importance of the Wiener filter in terms of hit rate alone.

Table 4.3 Picking accuracy

Routine time – Reference time	Calibration	Validation set	Complete set
Observations	1,213	226	1,439
Min (sec)	-1.533	-0.523	-1.533
Max (sec)	0.390	0.278	0.390
Mean (sec)	-0.037	-0.046	-0.038
STD (sec)	0.121	0.097	0.118

Automatic time (WF ¹) – Reference time	Calibration	Validation	Complete
Observations	1,232	226	1,458

Table 4.3 - continued

Min (sec)	-0.500	-0.217	-0.500
Max (sec)	0.800	0.501	0.800
Mean (sec)	0.007	0.022	0.009
STD (sec)	0.066	0.059	0.065
<hr/>			
Automatic time (UF ²) – Reference time	Calibration	Validation	Complete
Observations	1,098	183	1,281
Min (sec)	-0.500	-0.177	-0.500
Max (sec)	0.901	0.301	0.901
Mean (sec)	0.027	0.031	0.028
STD (sec)	0.092	0.055	0.087

¹ Wiener-Filtered.² Unfiltered.

The overall picking accuracy is summarized in Table 4.3. A clear pattern of the routine picking compared to my own picking is the tendency of the former to occur too early. Examples like the one displayed in Figure 4.3a abound and confirm the clear statistical tendency reported here. From the calibration set, the average discrepancy with the reference picking is -0.037 sec, with a standard deviation of 0.121 sec. The Wiener-filtered automatic picks are located much closer to the reference picks. The average difference is $+0.007$ sec, with a standard deviation of only 0.066 sec. The slightly positive average difference is a small residue from the delay so typical of the Baer-Kradolfer (1987) picking algorithm. The unfiltered automatic picking is less accurate, the average difference being $+0.027$ sec, with a standard deviation of 0.092 sec. Its overall accuracy is about intermediate between the routine picking and the Wiener-filtered automatic picking.

Figure 4.7 displays the discrepancies between the routine picking and the reference picking, and between the automatic picking on Wiener-filtered seismograms and the reference picking. The pattern of early picking (negative discrepancies) is clearly apparent for the routine picking, and discrepancies of -0.4 sec are observed even at SNR above $+30$ dB on some unfiltered seismograms. Since I picked or completely checked the picking and weighting of the calibration set 5 times without any interaction with a location algorithm, my conclusion is that systematic routine picking errors affect this data set. No such pattern is present in the automatic picking results, even if a few outliers are observed on the positive side. These few outliers result from a normal behavior of MannekenPix. Since the program applies a negative delay correction, I tuned it internally to preferably produce late arrivals than early arrivals for which the delay corrections would increase the picking error instead of reducing it. Tuning the picking algorithm to do so is quite easy since this is already its natural tendency (see 3.3.3). Figure 4.8 displays (-0.5 to $+0.5$

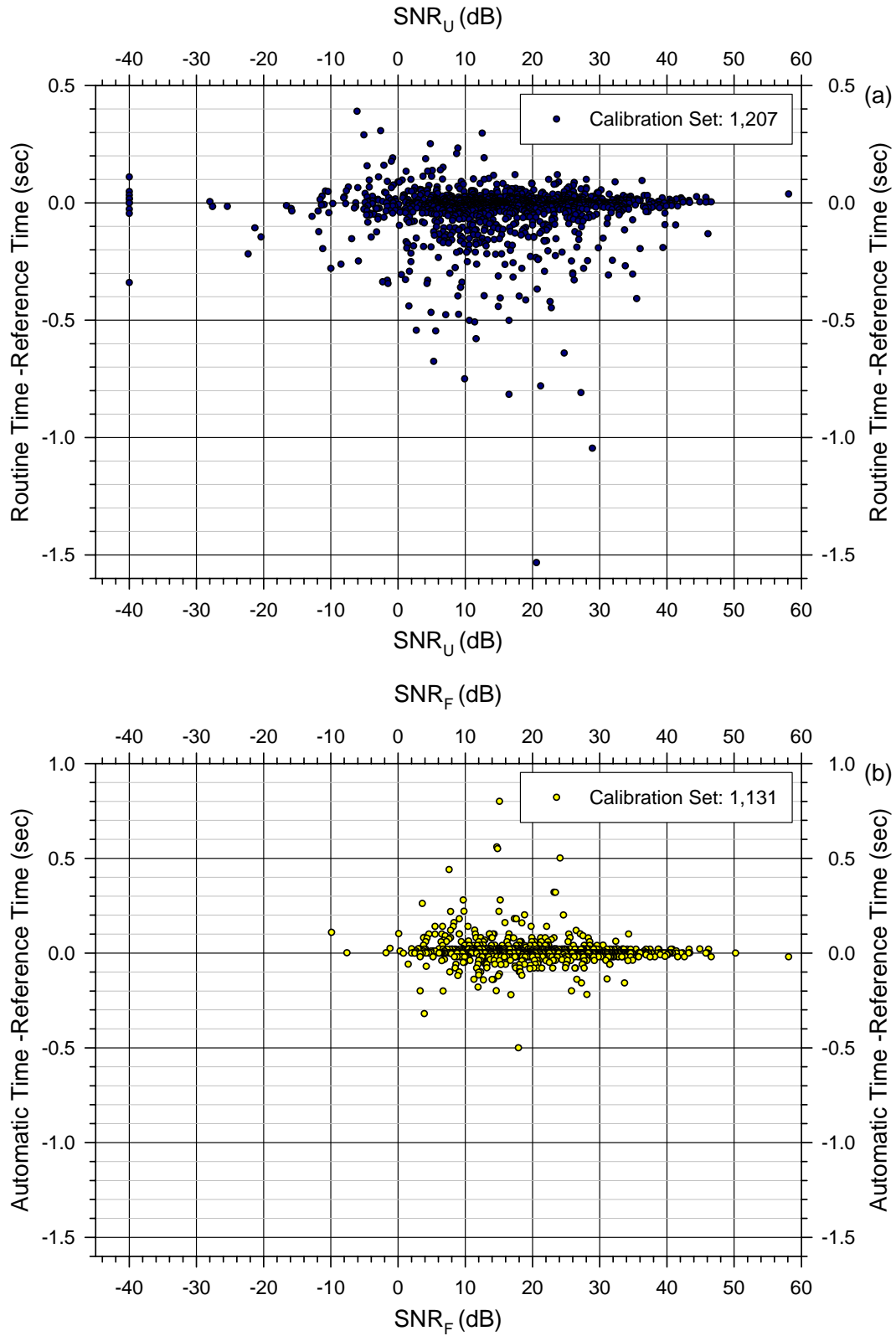


Figure 4.7 Estimated picking errors of the calibration set as a function of the SNR. (a) Routine picking, SNR from unfiltered seismograms around the routine pick time. (b) Automatic picking, SNR from Wiener-filtered seismograms around the routine pick time.

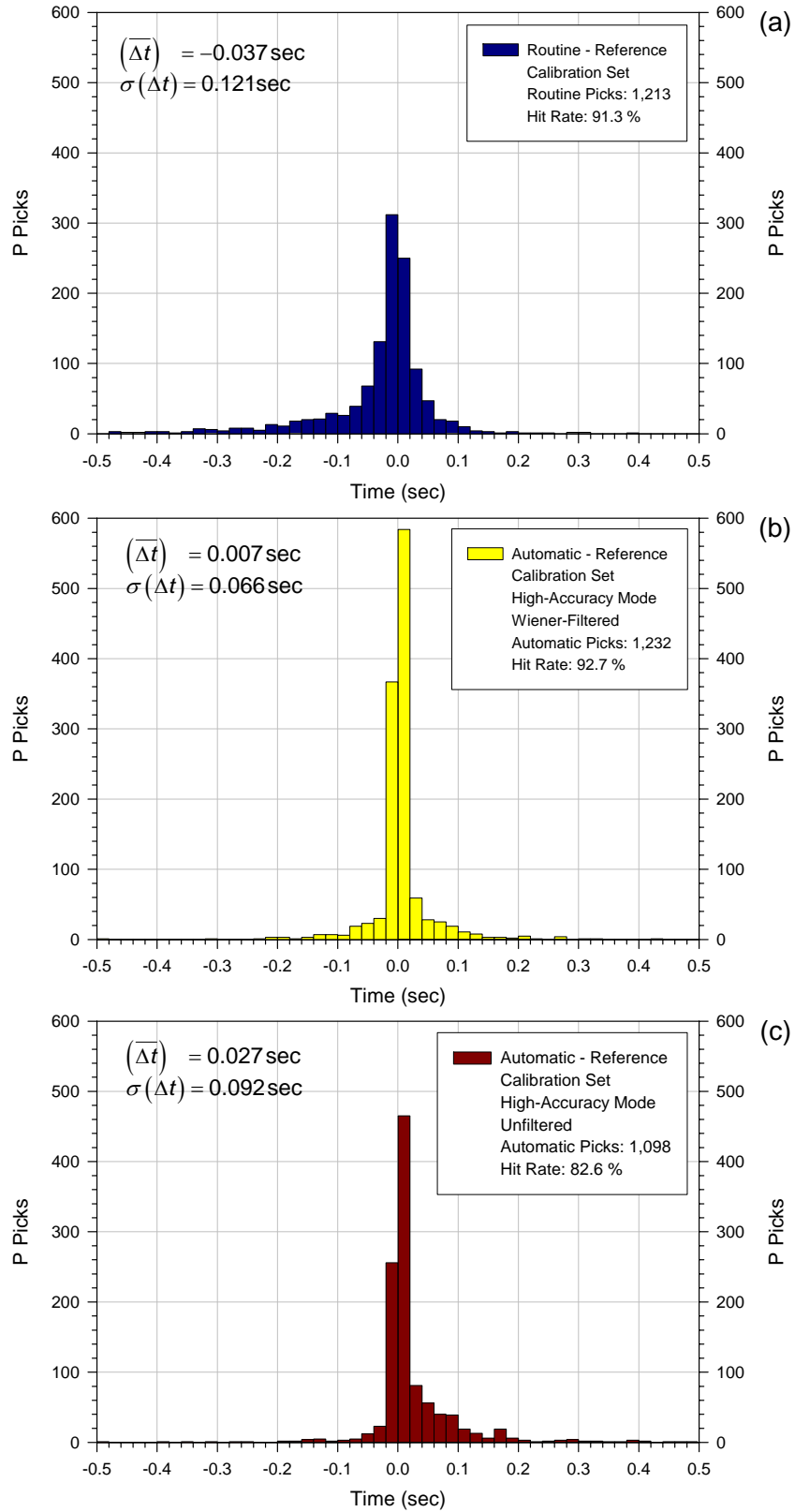


Figure 4.8 Estimated picking errors of the calibration set. (a) Routine picking. (b) Automatic picking, Wiener-Filtered. (c) Automatic picking, unfiltered.

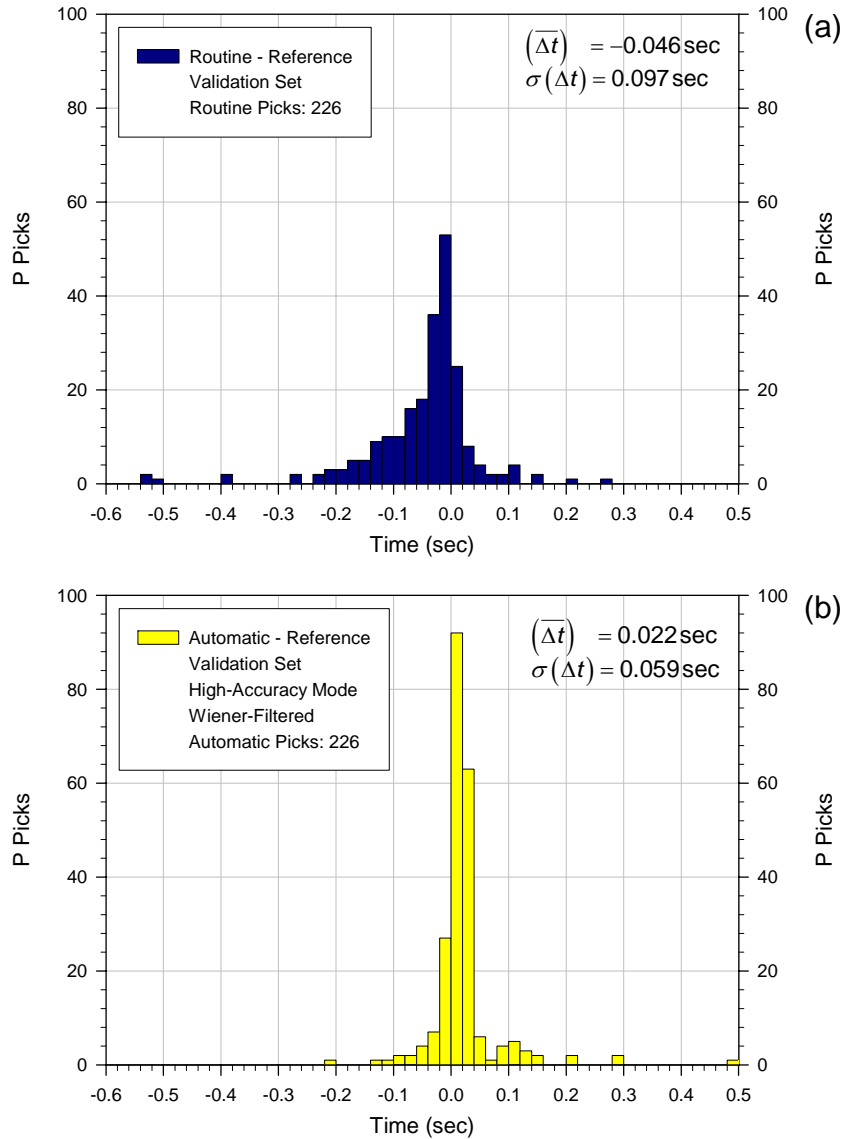


Figure 4.9 Estimated picking errors of the validation set. (a) Routine picking. (b) Automatic picking (Wiener-filtered).

sec) histograms of the estimated picking errors for the routine picking, and for the Wiener-filtered and unfiltered automatic pickings. Similar results are observed on the smaller validation set (Figure 4.9) from which the unfiltered automatic picking results have been omitted.

Since reference picks are only available for a limited number of seismograms, one way to collect information about the automatic picking accuracy for the complete data set is to compare the routine picks to the automatic picks. This is done in Figure 4.10a without the contribution of the calibration set, displayed in Figure 4.10b. If the automatic picks are more accurate than the

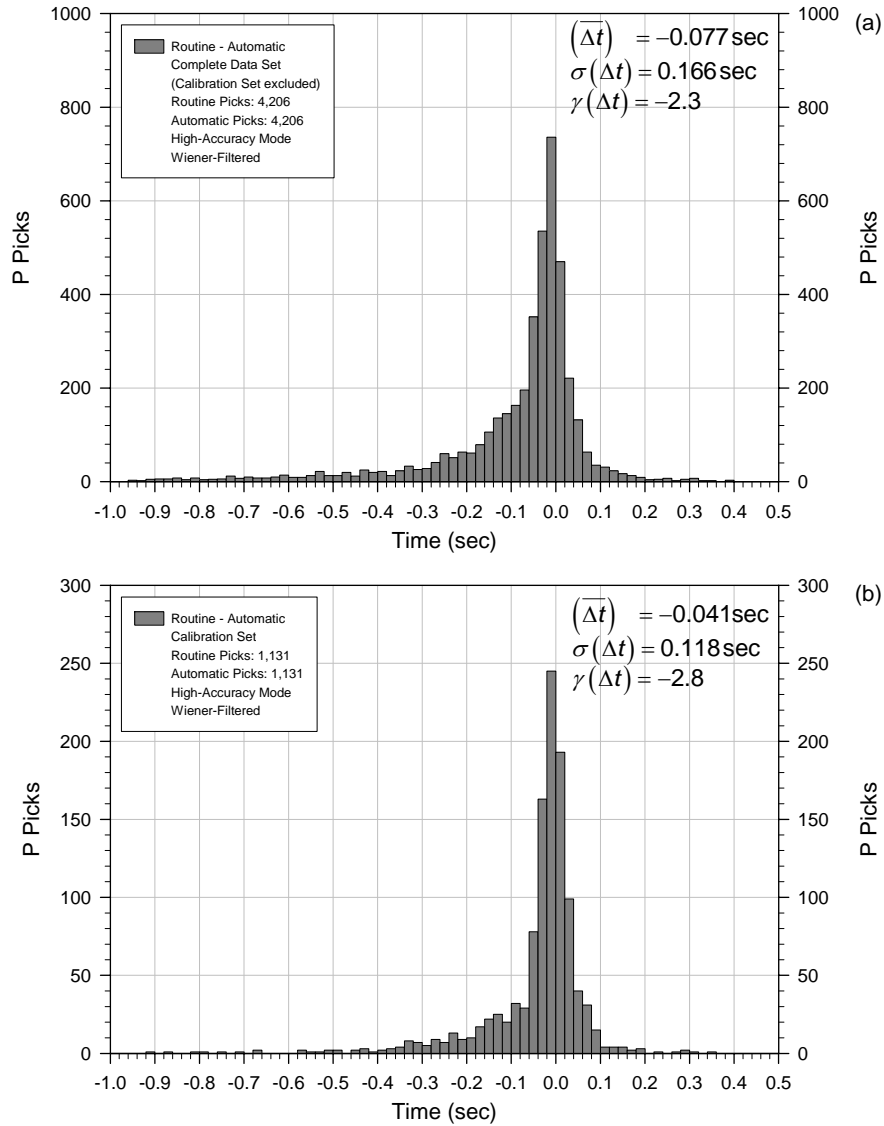


Figure 4.10 Difference between the routine pick times and the automatic pick times. (a) Complete data set, calibration set excluded. (b) Calibration set. γ is the skewness.

routine picks on the calibration set as previous results have shown, then it is very likely that they will also be more accurate on the complete data set since the distribution of discrepancies between automatic picks and routine picks is stable (Figure 4.10b and Figure 4.10a). This is not a proof in itself but it indicates that the overall pattern existing between the automatic picks and the routine picks is preserved when the picking is extended from the calibration set to the complete data set. If the automatic picking is more accurate than the routine picking on the calibration set, there is then no particular reason to believe that this might not be the case for the complete data set as well.

4.4 PICKING UNCERTAINTY

The uncertainties determined visually on the 1,329 reference picks of the calibration set are reported in Table 4.4. The target weight (TW) is the group membership that MannekenPix should be able to predict from the set of 9 predictors defined in 3.3.4.

Table 4.4 Weights of the reference picking (calibration set)

$0 \leq \varepsilon \leq 20$ ms	$20 < \varepsilon \leq 40$ ms	$40 < \varepsilon \leq 80$ ms	$80 < \varepsilon \leq 140$ ms	$ \varepsilon > 140$ ms	
PW 1	PW 2	PW 3	PW 4	PW 5	
255	857	199	18	0	
TW 1		TW 2		TW 3	
$0 \leq \varepsilon \leq 40$ ms		$40 < \varepsilon \leq 80$ ms		$80 < \varepsilon \leq 140$ ms	
1,112 (83.7 %)		199 (15.0 %)		18 (1.3 %)	
				0	

PW Picking Weight.

TW Target Weight.

ε is the estimated uncertainty on the reference picking.

The determination of the target weights for the automatic picks of the calibration set requires the definition of a rule derived from the automatic picking error Δt (Automatic time – Reference time) and from the uncertainty determined on the reference picks (the ε values in Table 4.1 and 4.4). The rule I adopted for the Dead Sea basin data is to select the target weight corresponding to the greatest value between the absolute automatic picking error $|\Delta t|$ and the absolute reference uncertainty $|\varepsilon|$. In other words, an automatic pick is given the same weight as the corresponding reference pick as long as it is located inside the uncertainty window of the reference pick. If it is located outside this uncertainty window, then the picking error prevails and is used to determine the target weight.

The target weights determined according to this rule for the 1,232 automatic picks of the calibration set are reported in Table 4.5. These target weights and corresponding predictors were imported into SPSS and a discriminant analysis was performed. Among the numerous outputs of such an analysis are Fischer’s linear discriminant coefficients that allow MannekenPix to make a prediction of the target weights on unseen cases. Only a very small selection of results is reported here.

Table 4.5 Target Weights of the automatic picking (calibration set)

TW 1	TW 2	TW 3	TW 4
$0 \leq \varepsilon \leq 40$ ms	$40 < \varepsilon \leq 80$ ms	$80 < \varepsilon \leq 140$ ms	$ \varepsilon > 140$ ms
952 (77.3%)	176 (14.3%)	65 (5.3%)	39 (3.1%)

TW Target Weight.

Box's M test examines the assumption of homogeneity of covariance matrices and it is also sensitive to multivariate normality. The test is significant on this data set and this means that covariance matrices differ between groups, violating one assumption of discriminant analysis. Discriminant analysis is however usually robust when the homogeneity of variances is not met.

An Anova test of equality of group means (Table 4.6) reveals that Wilks' Lambda is significant by the F test (at the 0.05 level) for all variables except GPDLDELFF. Since group means do not differ significantly for this variable, GPDLDELFF is probably a useless predictor on this data set.

Table 4.6 Tests of Equality of Group Means

Predictor	Wilks' Lambda	F	Degrees of Freedom	Significance	
WFILSTON	.918	36.427	3	1,228	.000
GPDLSTON	.811	95.211	3	1,228	.000
GPDLAMPR	.814	93.769	3	1,228	.000
GPDLSIGF	.923	33.927	3	1,228	.000
GPDLDELFF	.999	0.300	3	1,228	.825
THRCFRAT	.866	63.564	3	1,228	.000
PCABOTHR	.929	31.404	3	1,228	.000
PCBELTHR	.794	106.358	3	1,228	.000
CFNOIDEV	.943	24.965	3	1,228	.000

The standardized canonical discriminant function coefficients (Table 4.7) indicate the relative importance of the independent variables in predicting the dependent variable. Very small value for variable GPDLDELFF confirm the limited interest of this predictor. Table 4.7 suggests that GPDLSIGF might also be of little use on this data set.

Table 4.7 Standardized Canonical Discriminant Function Coefficients

Predictor	Function 1	Function 2	Function 3
WFILSTON	.239	-.038	.925
GPDLSTON	-.263	.677	-1.028
GPDLAMPR	-.361	-.520	1.879
GPDLSIGF	-.019	-.278	-.280
GPDLDELDF	.067	.064	.092
THRCFRAT	.055	.568	-1.209
PCABOTHR	.297	.864	.040
PCBELTHR	.651	.140	-.126
CFNOIDEV	.139	-.160	.731

The prediction results are reported in a confusion matrix (Table 4.8). Values on the main diagonal represent accurate predictions while off-diagonal values represent mis-classifications. A global rate of 66.8 % of cases is accurately classified.

Table 4.8 Confusion matrix (9 predictors)

		MPXW 1	MPXW 2	MPXW 3	MPXW 4	Total
Count	TW 1	708	115	109	20	952
	TW 2	35	79	48	14	176
	TW 3	18	17	22	8	65
	TW 4	7	10	8	14	39
<hr/>						
%	TW 1	74.4	12.1	11.4	2.1	100.0
	TW 2	19.9	44.9	27.3	8.0	100.0
	TW 3	27.7	26.2	33.8	12.3	100.0
	TW 4	17.9	25.6	20.5	35.9	100.0

66.8% of original grouped cases correctly classified.

TW Target Weight.

MPXW MannekenPix predicted Weight.

Variables GPDLDELDF and GPDLSIGF have then been excluded from the pool of predictors and a second analysis with the resulting 7 predictors has been performed. Original and cross-validated (jackknife) confusion matrices are reported in Table 4.9. The results are only

marginally better than with all 9 predictors. The cross-validation confusion matrix does not suggest that over-fitting might be an issue.

Table 4.9 Confusion matrices (7 predictors)

		MPXW 1	MPXW 2	MPXW 3	MPXW 4	Total	
Original	Count	TW 1	709	122	103	18	952
		TW 2	34	79	51	12	176
		TW 3	17	16	22	10	65
		TW 4	7	10	7	15	39
	Total		767 (62.3 %)	227 (18.3 %)	183 (14.9 %)	55 (4.5 %)	1232
	%	TW 1	74.5	12.8	10.8	1.9	100.0
		TW 2	19.3	44.9	29.0	6.8	100.0
		TW 3	26.2	24.6	33.8	15.4	100.0
		TW 4	17.9	25.6	17.9	38.5	100.0
Cross-validated	Count	TW 1	708	122	103	19	952
		TW 2	34	78	51	13	176
		TW 3	17	19	19	10	65
		TW 4	7	10	7	15	39
	Total		767 (67.0 %)	227 (18.3 %)	183 (14.9 %)	19 (1.6 %)	1232
	%	TW 1	74.4	12.8	10.8	2.0	100.0
		TW 2	19.3	44.3	29.0	7.4	100.0
		TW 3	26.2	29.2	29.2	15.4	100.0
		TW 4	17.9	25.6	17.9	38.5	100.0

67.0% of original grouped cases correctly classified.

66.6% of cross-validated grouped cases correctly classified.

In cross-validation, each case is classified by the functions derived from all cases other than that case.

Figure 4.11 is a visual display of the confusion matrix. Considerable smearing from the main diagonal values is observed for all target weights but the modal value of each row is always located on the main diagonal. This means that within each target weight group, the mode of

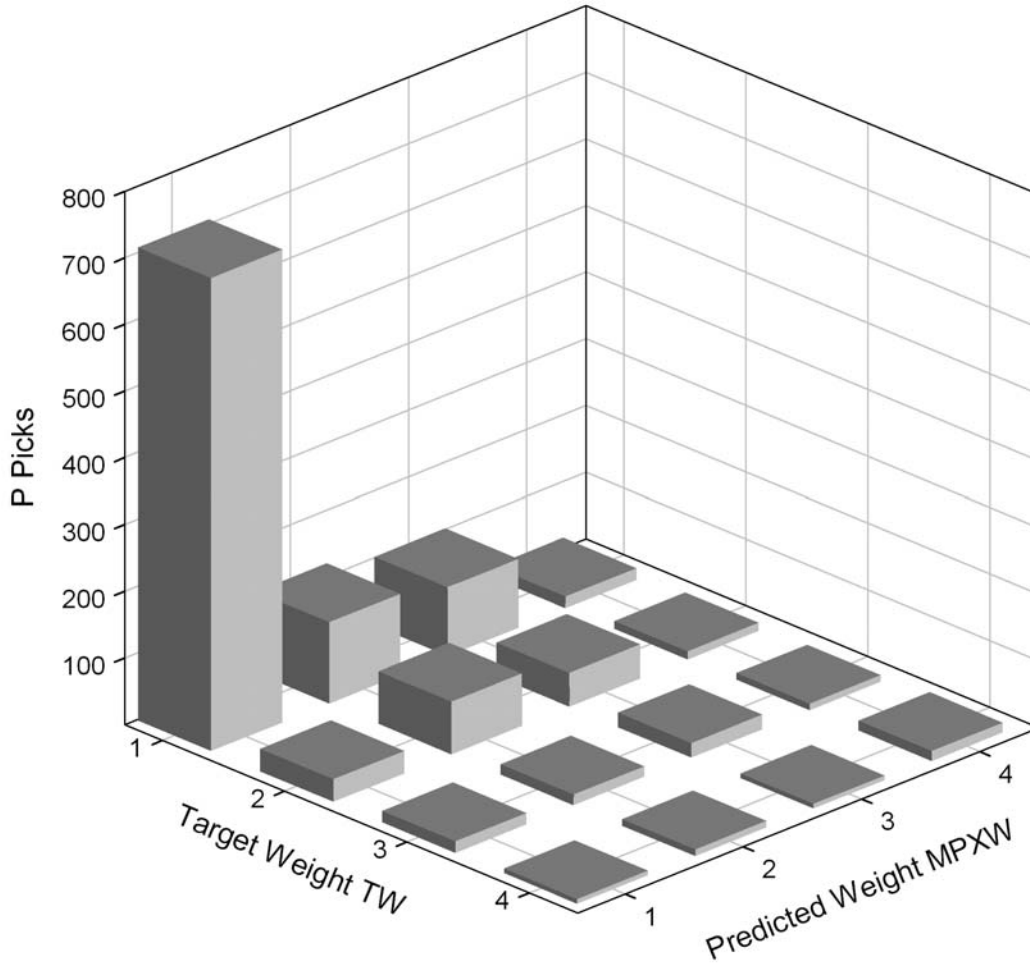


Figure 4.11 Confusion matrix (7 predictors).

classified cases occurs in the correct predicted weight. Of a higher interest to a seismologist is the content of each predicted weight group since this is what MannekenPix will provide for each automatic pick. This information is contained in the columns of the confusion matrix and it is specifically displayed in Figure 4.12. It is clear that target weight 1 data dominate predicted weights 1 to 3. Predicted weight 1 has however an excellent distribution with only minor contamination from target weights 3 and 4. Although dominated by the smearing from TW 1, predicted weight 2 data appear also to be rather good. The composition of predicted weight 3 differs little from the composition of predicted weight 2 but I would not consider this group as useless a priori. Predicted weights 4 are not dominated by the smearing of target weight 1 data and represent a small part (4.5 %) of the automatically picked calibration data that should not participate into any high-quality study. The automatic picking error content of the predicted weight groups is displayed in Figure 4.13. It shows that predicted weight 3 differs perhaps here

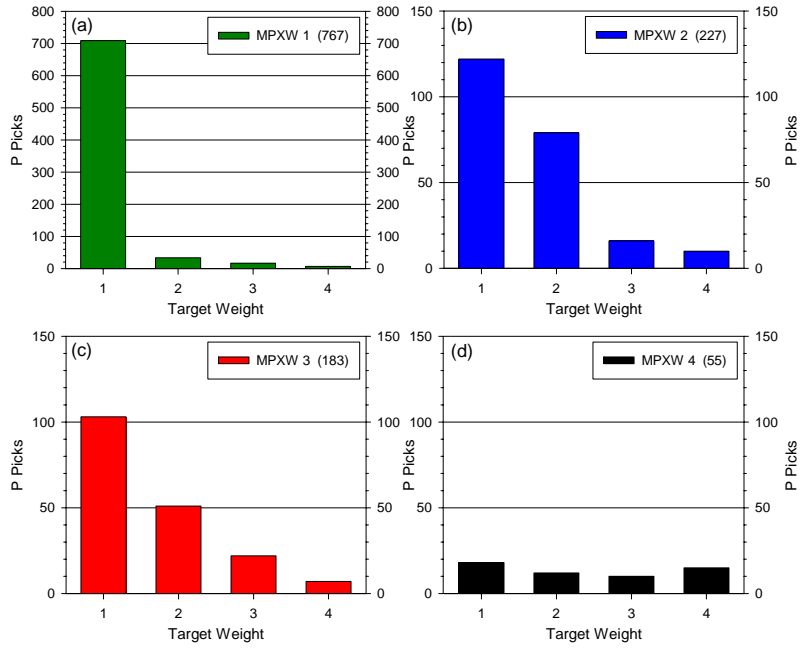


Figure 4.12 Distribution of Target Weights in the Predicted Weights.

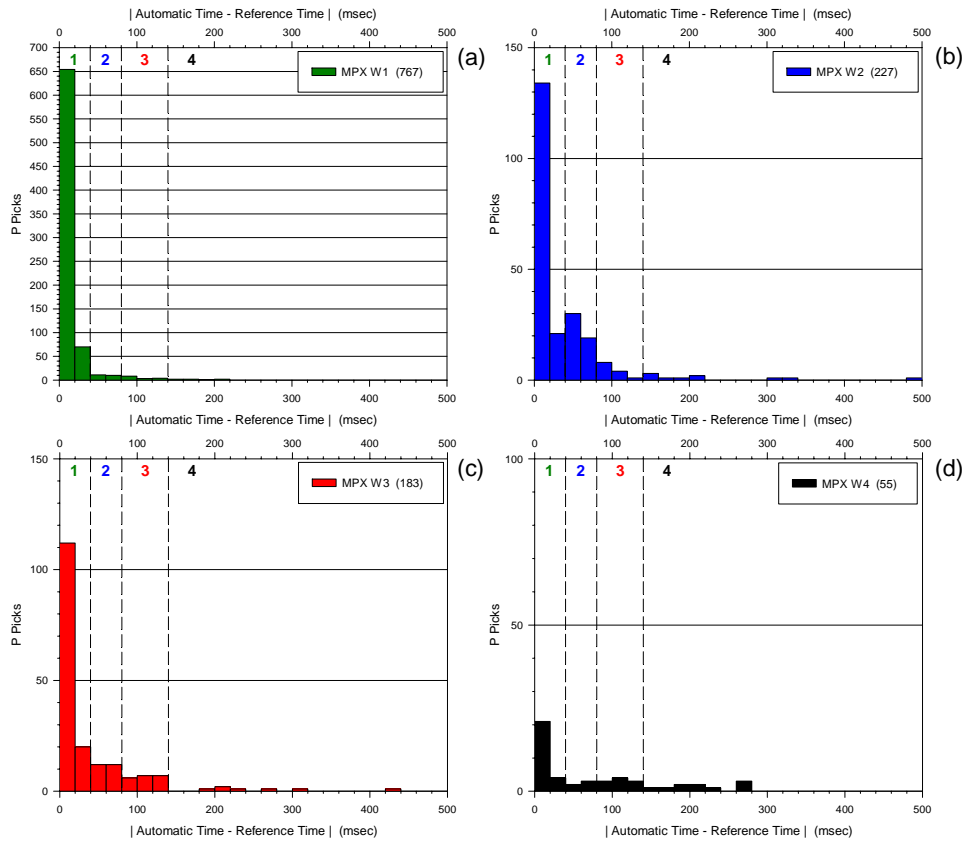


Figure 4.13 Distribution of automatic picking errors in the Predicted Weights.

slightly more from predicted weight 2 by the higher rate of picking errors correctly identified as belonging to predicted weight 3 group.

The confusion matrix from predictions made internally by MannekenPix on the validation set is reported in Table 4.10. No significant deviation is observed globally from the confusion matrix of the calibration set. An increase in the leakage from TW 1 toward MPXW 4 has however occurred.

Table 4.10 Confusion matrix of the validation set

		MPXW 1	MPXW 2	MPXW 3	MPXW 4	Total
Count	TW 1	126	22	21	11	180
	TW 2	2	13	5	2	22
	TW 3	3	2	8	3	16
	TW 4	0	1	3	4	8
Total		131 (57.8 %)	38 (16.8 %)	37 (16.4 %)	20 (8.9 %)	226
%	TW 1	70.0	12.2	11.7	6.1	100.0
	TW 2	9.1	59.1	22.7	9.1	100.0
	TW 3	18.8	12.4	50.0	18.8	100.0
	TW 4	0	12.5	37.5	50.0	100.0

66.8 % of original grouped cases correctly classified.

TW Target Weight.

The picking results for the complete data set are reported in Table 4.11. The complete data set exhibits a significantly reduced rate of predicted weights 1 compared to the calibration set, about the same rate of weights 2 and 3 as in the calibration set, and a much higher rate of weight 4 picks. Although a higher leakage from TW 1 to MPXW 4 was observed on the validation set, it does not explain the significantly reduced rate of MPXW 1 and the increase in the number of MPXW 4 values. The complete data set is more difficult to pick reliably for MannekenPix than the calibration set.

Table 4.11 Predicted Weights of the automatic picking (complete data set)

MPXW 1	MPXW 2	MPXW 3	MPXW 4
3,058 (43.1 %)	1,173 (16.5 %)	1,554 (21.9 %)	1,304 (18.4 %)

4.5 A SIMPLE RELOCATION

A relocation of earthquakes from the calibration set has been performed using successively the routine picks, the reference picks, and the automatic picks. The purpose of this simple relocation is to see if some reduction in the RMS residual times is observed by re-picking the calibration set. The original analyst weights have been applied for the routine picks, the target weights (TW) have been applied for the reference picks and the predicted weights (MPXW) have been applied for the automatic picks. Model Israel (see 2.3) was used and the relocations were performed with Velest.

Table 4.12 Locatable earthquakes of the calibration set

Picking Type	Earthquakes	Locatable	Not locatable
Routine		110 (96.5 %)	4 (3.5 %)
Reference	114	110 (96.5 %)	4 (3.5 %)
Automatic		109 (95.6 %)	5 (4.4 %)

The number of locatable earthquakes is reported in Table 4.12 and does not differ significantly between the three picking groups. The distribution of RMS residual times for the three relocations is displayed in Figure 4.14.

The most important result is that the reference and automatic pickings display higher residual times than the routine picking does, especially between 0.1 and 0.3 sec. This could imply that the routine picking is more accurate as a whole than the reference picking. It can also imply that the reference picking is more accurate than the routine picking but that the velocity structure of the region imposes high residuals on earthquake locations derived from a simple 1-D model without station corrections. Since we know that the Dead Sea region has a complex velocity structure (see Chapter 2), higher residuals are not incompatible with a greater picking accuracy. In any case, the convoluted interactions between arrival times, velocities and source locations can only be properly understood if the picking is done objectively and is free from the attractive influence of location results. The question of residual times appears to be a delicate issue here and it should best remain open until a high-quality tomographic study including all relevant variables will shed the appropriate light on the problem.

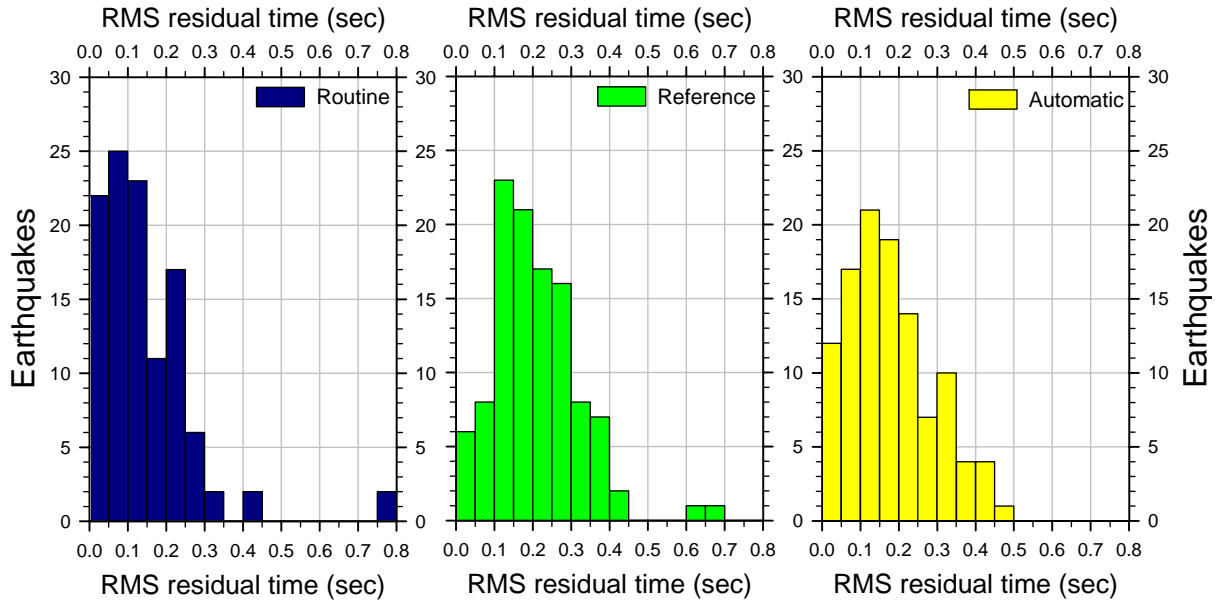


Figure 4.14 Relocation residual times (calibration set)

The number of locatable earthquakes for the complete data set is reported in Table 4.13. The number of locatable earthquakes derived from the automatic picking is only very slightly lower than the number of locatable earthquakes derived from the routine picking.

Table 4.13 Locatable earthquakes of the complete data set

Picking Type	Earthquakes	Locatable	Not locatable
Routine	531	528 (99.4 %)	3 (0.6 %)
Automatic		526 (99.1 %)	5 (0.9 %)

Chapter 5

Conclusions

Lower-crustal strength under the Dead Sea basin

Conclusions about the lower-crustal seismicity of the Dead Sea basin were drawn in section 2.5 (pp. 19-21). Furthermore, results from Ginzburg and Ben-Avraham (1997) show that a major basinal fault affecting the crystalline basement at 12-15 km depth runs exactly at the location ($X = 190$, $Y = 80$ in Figure 2.4a) where a swarm of micro-earthquakes nucleate in the lower crust below 25 km. This evidence suggests that major faulting in the Dead Sea basin extends probably deeply into the lower crust. There is little doubt then that a joint study combining seismic reflection profiles and earthquake focal mechanisms should be made at this location in an attempt to relate the lower-crustal seismicity to the more superficial and better documented tectonics. By such a study, further insight might be gained into the role played by lower-crustal brittle failure in the development of the basin.

The rheological stratification of the lithosphere in the Dead Sea region should also be studied further, preferably by several independent methods. For instance, the effective elastic plate thickness could be derived from gravity data. On the other hand, postseismic relaxation studies could also be performed, but as these studies depend critically on the occurrence of moderate to large earthquakes (with long recurrence intervals along the DST), an unknown waiting time is required. One study, from GPS data spanning only a short time window, has been performed in the Middle-East (Piersanti et al., 2001) following the $M_w = 7.2$ Gulf of Aqaba-Elat earthquake of November 1995. The best agreement with the data was obtained with a low viscosity value (10^{18} Pa s) for the lower crust and the asthenosphere, a result consistent with the higher heat flow (see section 2.5) in the Gulf of Aqaba-Elat compared to the heat flow in the Dead Sea basin.

Following analytical and numerical models recently developed, Roy and Royden (2000a, 2000b) suggest also that “when an elastic upper crust is underlain by a low-viscosity lower crust at a strike-slip plate boundary, the deformation zone tends to be broad at the surface and encompasses many parallel strike-slip faults in an interacting network. In contrast, when the entire crust behaves elastically, the deformation zone remains narrow and focused on a single plate-bounding fault, reflecting imposed mantle motions”. It would be interesting to apply these models to the DST in the vicinity of the Dead Sea basin, and to compare the results with those obtained for the Gulf of Aqaba-Elat. The relative simplicity of the DST in the vicinity of the Dead Sea basin might well be closely related to the presence of a strong lower crust along this segment of the plate boundary.

Several other methods can be used to better constrain the rheological properties of the lower crust and the upper mantle, and new techniques are regularly proposed. The Dead Sea region gives us a rare opportunity to better understand the properties of the tectonically active continental lithosphere under a low heat flow regime. It would however not be wise to possibly miss such an opportunity by reaching conclusions too soon about the rheological properties of the uppermost mantle, for which we only know that it appears to be aseismic during 14 years. This is especially true when taking into account the complexity of the problems involved and the erroneous generalizations already made during the last 20 years regarding the rheology of the continental lithosphere.

A new picking system: MannekenPix

The development of MannekenPix convinced me that many picking algorithms are not routinely used on a large scale because the complementary components they need are just missing. Designing a turn-key automatic picking package required me to work on a broad front of problems, sometimes far from the central theme of automatic picking. An adaptive Wiener filter, delay corrections and a sophisticated weighting engine were necessary ingredients to turn the Baer-Kradolfer (1987) picking algorithm into a real production tool compatible with the stringent requirements of high-quality studies. MannekenPix is still far from being an ideal automatic picking system but I believe that as it is now, it may already benefit to a significant number of applications ranging from automatic bulletin generation to automatic focal mechanism determinations and high-quality tomographic studies.

For data sets in which the dominant frequencies of signal and noise are sufficiently distinct, the automatic Wiener filter of MannekenPix can significantly increase the SNR of noisy seismograms. Such a filter does neither affect significantly the amplitude of the signal nor does it introduce delays or precursory sidelobes like sub-optimal filters generally do. For data sets in which the dominant frequencies of signal and noise mainly overlap and no improvement by traditional filtering is expected, the Wiener Filter of MannekenPix just leaves the original data intact. The filter is probably too slow to be integrated as is in a near real-time picking system, but it can run significantly faster with a minor trade-off in the search of the optimal SNR.

Another quality of MannekenPix is that its picking weights derive from true uncertainty classes defined in time units. The weighting mechanism of MannekenPix attempts to predict the same uncertainties as those that would be estimated by the analyst. As such, it acts as an extension of manual weighting and not as a totally independent method possibly in conflict with the legacy. Since MannekenPix weights can be computed very fast, any picking system (real-time or not) could benefit from the method. However, the discrimination of quantitative time uncertainty groups depends not only on the MannekenPix method of weighting but also on the availability of adequate predicting variables. There is thus no guarantee that predictors of an existing picking system can provide such discriminating power.

The long-term goal of automatic weighting should be perhaps to determine picking uncertainties without any user-supplied reference weights. By doing so, another source of inconsistencies might be removed from automatic picking data sets. One possible way to reach this difficult goal requires however that the deconvolution by the seismograph response would be both routinely feasible and useful within the framework of automatic picking.

The main strength of MannekenPix is probably its ability to provide a good rate of accurate picks. About its limitations in the actual state, MannekenPix misses some arrivals that it should not miss. The difficult problem of quantitatively weighting the picks needs also to be more reliable, but the highest weight classes are however well insulated from largely erroneous picks.

Deconvolution

The main problem with the recording system is that it usually acts as an out-of-focus lens through which everything appears blurred. Such a distortion implies a systematic reduction of the signal-to-noise ratio, even in the absence of significant additive noise. Refocusing the signal by

the digital deconvolution of the instrumental response could significantly benefit to the accurate picking and autonomous weighting of phases, but this possibility needs to be fully evaluated with real-world data. From a practical standpoint, deconvolution should be largely feasible since the instrumental response is usually well known for most data sets. This response may not be perfectly stable over time but it is probably stable enough for the purpose of deconvolution. It would also be interesting to investigate whether or not deconvolution has an incidence on the detection rate of weak signals buried in noise.

Application of MannekenPix to the Dead Sea region

The application of MannekenPix to the Dead Sea local data convinced me also that the program is able to detect and correct a significant number of systematic errors commonly present in data sets. These errors do not only include large picking errors, they also include largely incorrect observation weights. Naturally, the program also makes similar errors but in a limited way that can be clearly assessed during the calibration of the weighting engine.

By using MannekenPix to pick the seismograms from well-constrained earthquakes of the Dead Sea region, 526 (99.1 %) out of 531 earthquakes are locatable by automatic picking. Out of 15,250 seismograms, 6,889 (45.2 %) were routinely picked and 7,089 (46.5 %) could be picked automatically. On the calibration data set, the average discrepancy between the routine picking and the reference picking is -0.037 sec, with a standard deviation of 0.121 sec. The average discrepancy of the automatic picking is $+0.007$ sec with a standard deviation of only 0.066 sec. A clear pattern of early picking affects the routine picking of this data set while the automatic picking discrepancies are significantly reduced and nearly randomly organized compared to those of the routine picking.

For the complete data set, 3,058 phases (43.1 %) of the 7,089 automatic P picks fall into predicted weight 1 (absolute uncertainty not greater than 40 msec), 1,173 (16.5 %) fall into class 2 (absolute uncertainty between 40 msec and 80 msec), 1,554 (21.9 %) fall into class 3 (absolute uncertainty between 80 msec and 140 msec) and 1,304 (18.4 %) fall into class 4 (absolute uncertainty greater than 140 msec). The results from a totally out-of-sample set of reference picks and weights confirm the stability of predicted weights on unseen data.

The automatic filtering and picking of the complete data set with MannekenPix 1.7.6 took on the average 0.4 sec per seismogram (2.5 seismograms / sec, or 9000 seismograms / hour) with a 1.4 GHz Pentium 4 processor, and Linux as operating system.

My understanding is that MannekenPix fulfilled the main purpose of this work, i.e. to provide a set of picking data satisfying the quantity and quality requirements for a travel-time tomographic study of the Dead Sea region.

Application of MannekenPix to the seismicity of Italy

Di Stefano et al. (2004, to be submitted) present the application of MannekenPix to about 240,000 seismograms from nearly 29,000 local earthquakes routinely recorded by the Italian national seismic network during the period 1998-2001. A standard version of MannekenPix was able to produce 103,131 P picks (73% of the 139,500 routine onsets) from 23,108 events out of 28,900 events.

While the bulletin readings are almost always un-weighted, MannekenPix distinguishes between higher and lower quality picks. About 17,130 phases (17 %) out of the 103,131 fall into class 1 (absolute uncertainty not greater than 0.1 sec) and 15,429 phases (15 %) fall into class 2 (absolute uncertainty between 0.1 sec and 0.2 sec). These figures are slightly lower than for the 700 seismograms of the calibration set, possibly because a greater number of events of small magnitude is present in the complete data set, yielding a much higher number of phases in class 4.

To test the re-picked data set we merged the 103,131 automatic picks with bulletin phases from other networks included in the CSI bulletin (Catalogo della Sismicita' Italiana) and we relocated the events following the same method and parameters used to generate the catalog. We replaced all the INGV station bulletin readings with the automatic picks. Results displayed in Figure 5.1 show that MannekenPix arrival times for classes 1 and 2 combined, produce a distribution of residual times with a much smaller standard deviation than for the CSI bulletin. This means that a significant increase in the quality of the data set has been gained, coming at the price of a reduced quantity. It is important to note that these results were obtained with MannekenPix 1.6.2 that does not include yet the secondary delay correction. The polarities provided by MannekenPix make also the subset suitable for high-quality automatic focal mechanism determinations on a large scale with respect to the bulletin.

The application of MannekenPix to the very large and inhomogeneous dataset of waveforms recorded in Italy from 1998 to 2001 yields nearly 33,000 high-quality weighted picks and polarities. By using MannekenPix we achieved our goal to build a new dataset of P-wave arrival times and polarities belonging to user-defined classes 1 and 2 (the highest qualities) with associated uncertainty estimates. The analysis of relocation residuals and the time versus distance plots for these classes show the effectiveness of the picking system. Although the hit rate of MannekenPix 1.6.2 applied to this large and noisy dataset did not exceed 75 % of the rate of an expert seismologist, the consistency of the automatic arrivals and polarities and their rapid estimation are very suitable to substitute bulletin data, solving the typical problem of extending consistency and quality to large datasets. We believe that the improved seismic dataset is suitable to refine the seismic tomography, and to improve the knowledge of the local and regional stress fields in the Italian peninsula through the determination of a large number of focal mechanisms derived from first-motion polarities determined automatically by MPX for well-constrained events, and extending also the analysis to small magnitudes and past events.

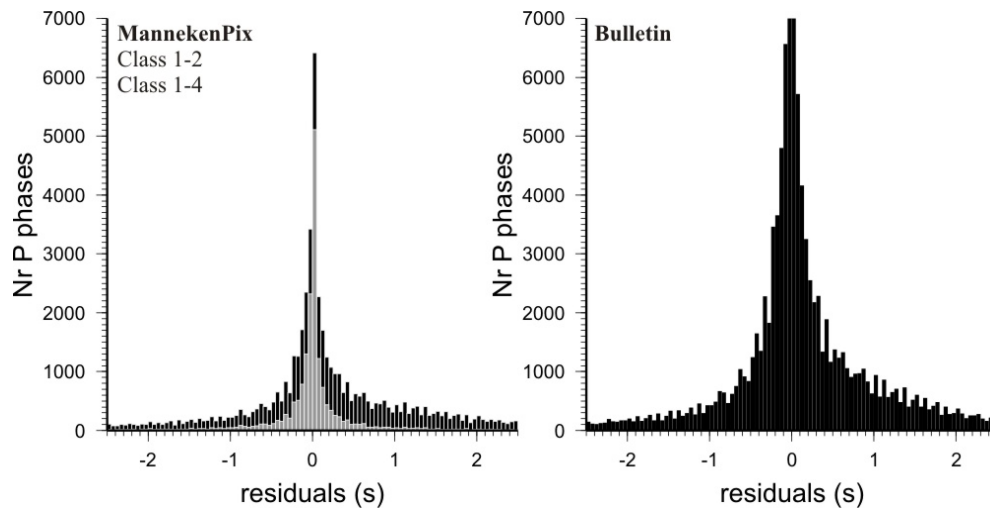


Figure 5.1 Distribution of relocation residual times for INGV stations. (Left) MPX automatic picking residual times. Classes 1-4 in black and classes 1-2 in gray. (Right) CSI bulletin residual times.

Appendix A

Abbreviations

AIC	An Information theoretic Criterion
AR	Autoregressive
CSI	Catalogo della Sismicita' Italiana
dB	decibel
DFT	Discrete Fourier Transform
DPSS	Discrete Prolate Spheroidal Sequence
DST	Dead Sea Transform
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FPE	Final Prediction Error
GII	Geophysical Institute of Israel
GPS	Global Positioning System
INGV	Istituto Nazionale di Geofisica e Vulcanologia
JSO	Jordan Seismological Observatory
LET	Local Earthquake Tomography
LTA	Long-Term Average
MDA	Multiple Discriminant Analysis
MEM	Maximum Entropy Method
MPXW	MannekenPix Predicted Weight
PSD	Power Spectral Density
PW	Picking Weight
RMS	Root Mean Square
SMA	Simple Moving Average
SMSTD	Simple Moving Standard Deviation

SNR	Signal-To-Noise Ratio
STA	Short-Term Average
STDB	Simple Standard Deviation Band
TW	Target Weight
WSS	Wide-Sense Stationarity

Appendix B

Source function parameters and seismograph response of the Wiener-filtered synthetic example (Figure 3.6)

The displacement source function used to build the synthetic example of Figure 3.6 is plotted in Figure B.1a. It was generated in Pitsa (Scherbaum and Johnson, 1992) according to Brune's model (Brune, 1970) but all other operations were performed in Matlab, mostly with functions of the Signal Processing Toolbox. Stress drop, shear velocity and source radius have typical values for a microearthquake of magnitude $M_L = 2$ in the Dead Sea region (van Eck and Hofstetter, 1989). The value of the shear modulus corresponds to a source depth of 5-15 km (Turcotte and Schubert, 2002). The displacement function has been truncated to a length of 200 msec, and the corresponding velocity function is plotted in Figure B.1b. The sampling period is 10 msec throughout the example. In order to build a P-wave pseudo-wavetrain, a series of spikes of decreasing amplitude (Figure B.1c) was generated starting at time $t = 7$ sec of a 10 seconds length simulation segment. The interval of time between spikes is 200 msec and equals the duration of the truncated source functions.

The source velocity function was convolved by the spikes to produce the noiseless pseudo-wavetrain of Figure B.2a. Although the first spike is located exactly at time $t = 7$ sec, the wavetrain at that time has a zero value. This results from the fact that the first sample of the source velocity function has a zero value and that this value is preserved through the convolution by a spike. By analogy of continuous functions, I consider the first-arrival time of a noiseless wavetrain to be the last zero-valued sample of the pre-arrival segment and not the first sample different from zero. This choice looks natural when seismograms are visualized in the perspective of onset picking.

In order to add noise to the pseudo-wavetrain, 10 seconds of instrumentally recorded noise were deconvolved by the response of a digital seismograph composed of a simulated Marks L4-C

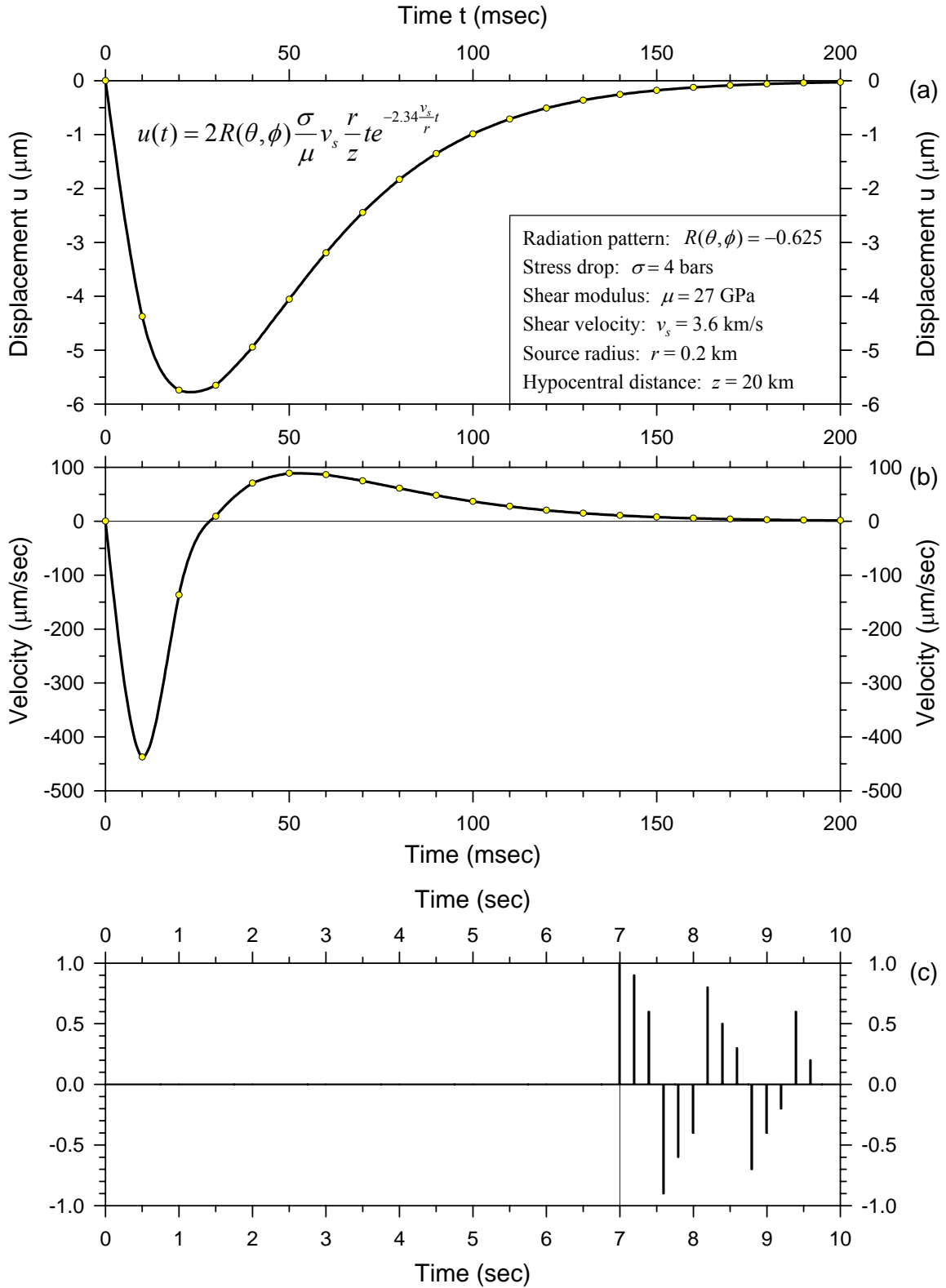


Figure B.1 Synthetic source function and convolution spikes. (a) Source function (displacement). (b) Source function (velocity). (c) Convolution spikes. The time of first arrival is marked by the dark red line at time $t = 7$ sec. The sampling period is 10 msec.

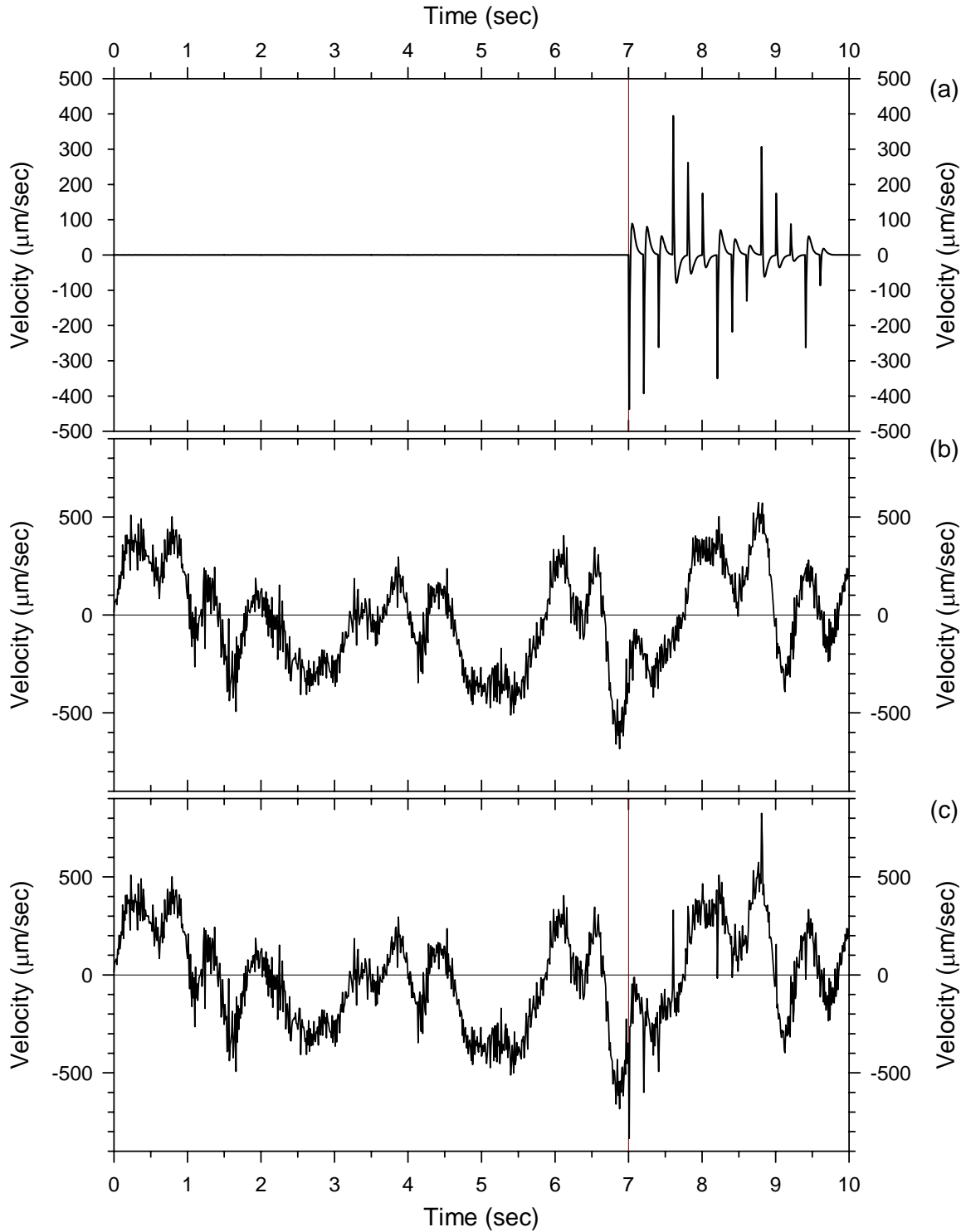


Figure B.2 Velocity signal, noise and noisy signal. (a) Noiseless synthetic pseudo-wavetrain signal (velocity). (b) Deconvolved recorded seismic noise. (c) Combined signal (a) and noise (b). The time of first arrival is marked by the dark red line at time $t = 7$ sec. The sampling period is 10 msec.

velocity seismometer and a digital 5th-order Butterworth low-pass (anti-aliasing) filter. The response of this digital simulated seismograph is close to the response of the actual instruments deployed by GII in the Dead Sea region (van Eck and Hofstetter, 1989), except for a high-pass filter with corner frequency at 0.2 Hz that has been discarded here. The deconvolution was performed by spectral division in Matlab with routine “Decon1” of seismological package Coral¹ (Creager, 1997). The primary purpose of Decon1 is to compute receiver functions (Langston, 1979) but nothing prevents other uses like the one described here. Deconvolving the recorded noise by the seismograph response serves to better simulate true ground noise, the noise that mixes with the signal before both are filtered by the seismograph during the recording process. The result is plotted in Figure B.2b. The noiseless signal (Figure B.2a) and the deconvolved noise (Figure B.2b) were then added together. The result is the noisy signal plotted in Figure B.2c. All plots in Figure B.2 are supposed to be free of instrumental distortions and are expressed as ground velocity.

The seismograph response has already been used to deconvolve the instrumentally recorded noise segment but no details were provided about how this response was obtained.

The discrete velocity impulse response of a standard Marks L-4C seismometer can be determined (Scherbaum, 1996) from the analog zeros and poles provided by the manufacturer (Table B.1). I did this in Matlab by computing the transfer function $T(s)$ from the analog zeros and poles with function “zp2tf”, and by mapping $T(s)$ onto the discrete transfer function $T(z)$ by the bilinear transformation (function “bilinear” in Matlab). An arbitrary generator constant $G = 1.0$ [V/(m/sec)] was used. The discrete impulse response and frequency response are then given by functions “impz” and “freqz” respectively. The results are plotted in Figure B.3.

Table B.1 Zeros and Poles of the Marks L-4C

Zero	Pole
0 + 0 i	- 4.443 + 4.443 i
0 + 0 i	- 4.443 - 4.443 i

Eigenfrequency $f_c = 1.0$ Hz , Damping $h = 0.707$.

¹ <http://www.ess.washington.edu/Faculty/creager/>

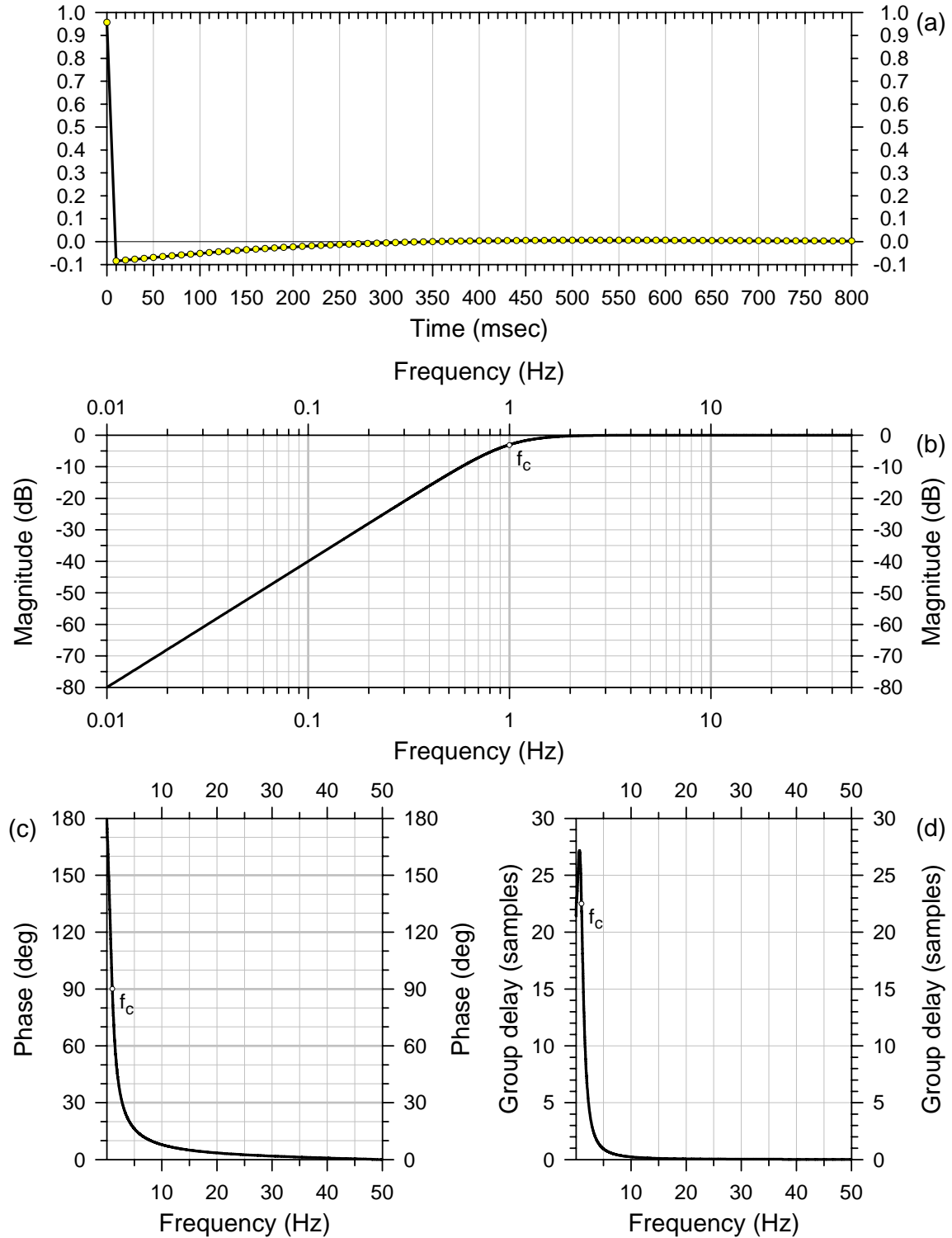


Figure B.3 Digital simulated L4-C velocity seismometer. (a) Velocity impulse response. (b) Magnitude spectrum. (c) Phase spectrum. (d) Group delay spectrum. The sampling period is 10 msec.

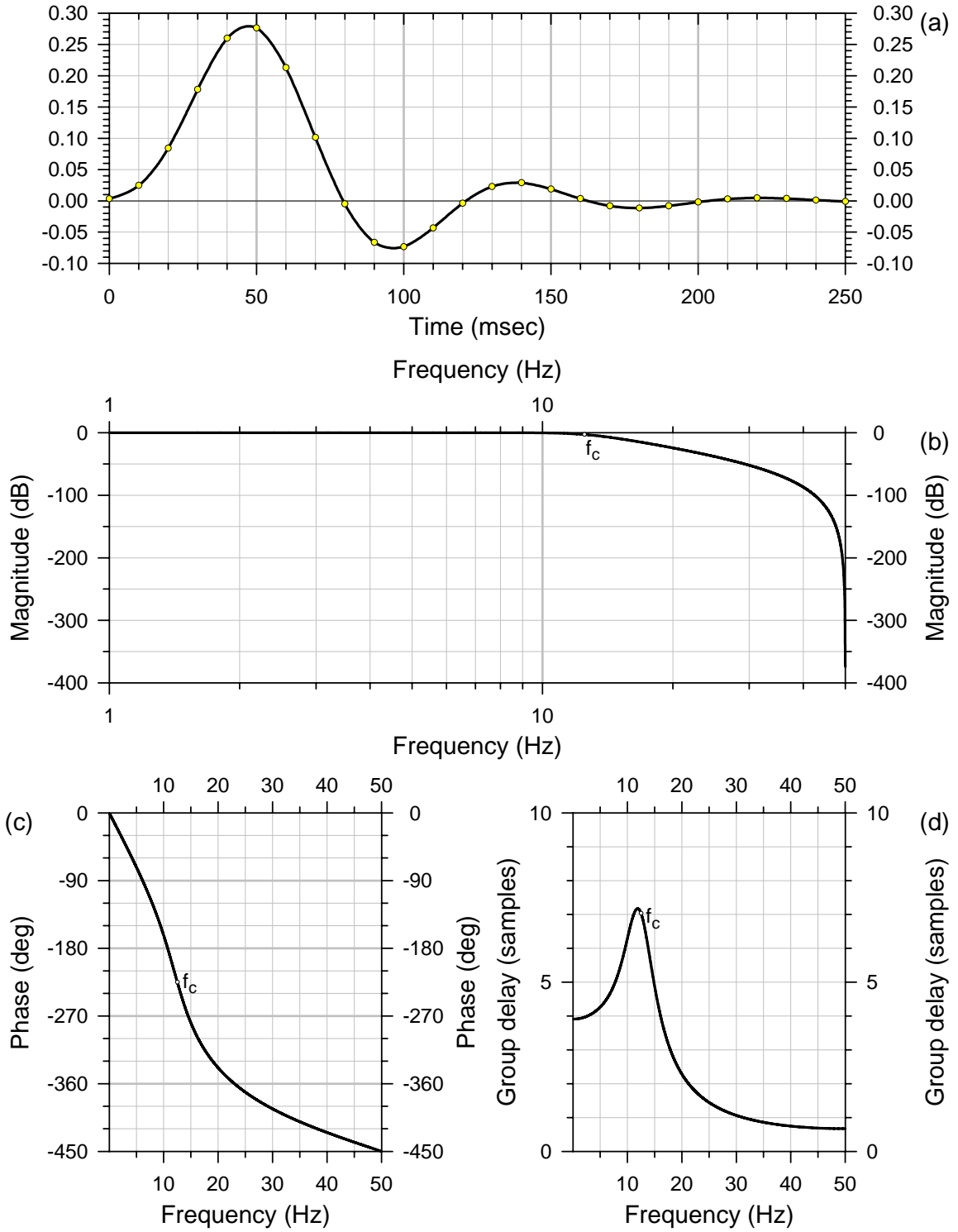


Figure B.4 Digital 5th-order Butterworth anti-aliasing filter. (a) Impulse response. (b) Magnitude spectrum. (c) Phase spectrum. (d) Group delay spectrum. The sampling period is 10 msec.

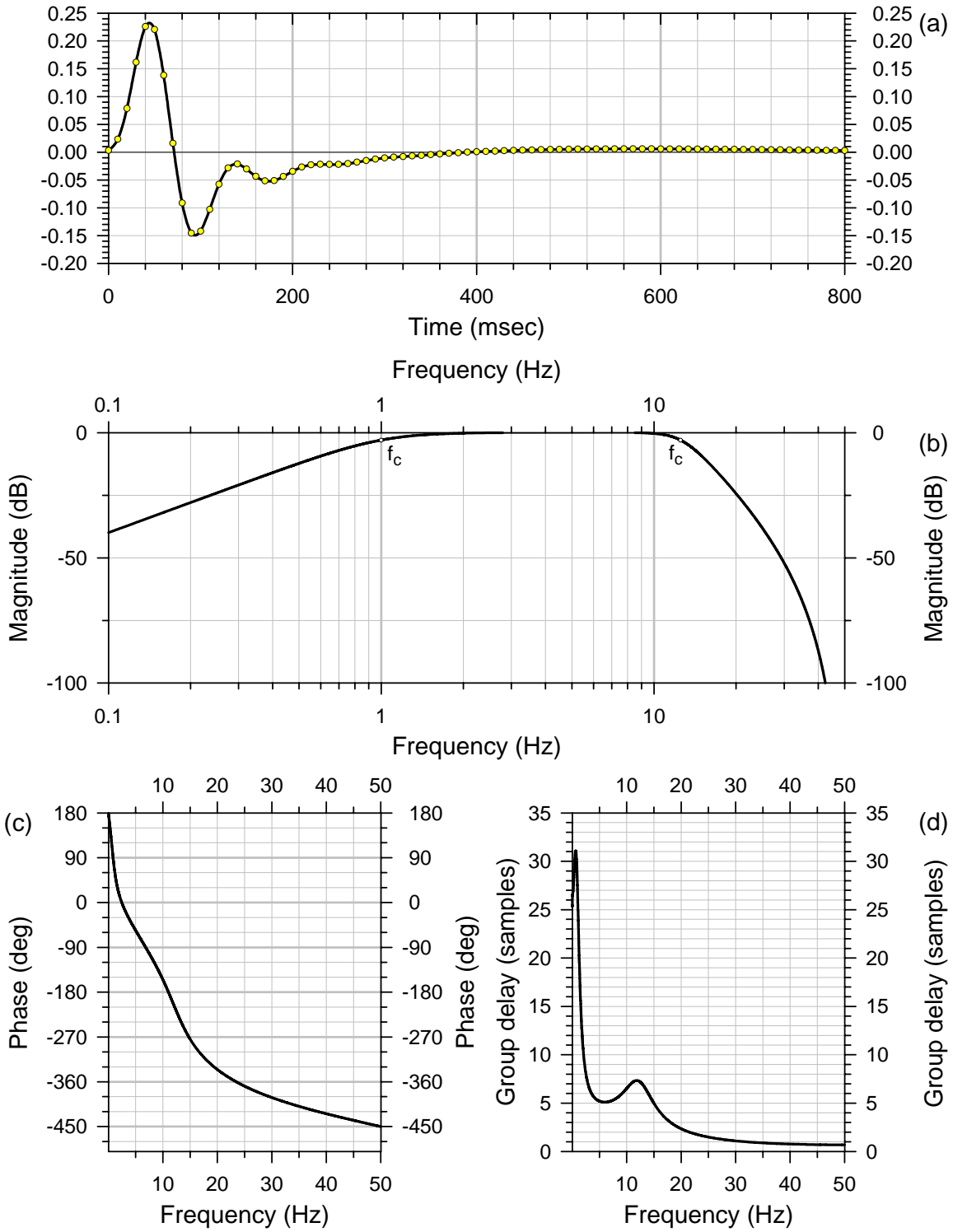


Figure B.5 Simulated seismograph. (a) Impulse response. (b) Magnitude spectrum. (c) Phase spectrum. (d) Group delay spectrum. The sampling period is 10 msec.

The 5th-order Butterworth digital anti-aliasing filter with corner frequency $f_c = 12.5$ Hz was computed with function “butter” of Matlab. This function creates a Butterworth analog low-pass filter, and converts this analog filter into a digital filter through a bilinear transformation with warped corner frequency. The discrete impulse response and frequency response are then given by functions “impz” and “freqz” as for the seismometer. The results are plotted in Figure B.4.

The seismograph response (combined L4-C and anti-aliasing filter) is plotted in Figure B.5. The simulated seismogram of Figure 3.6a was finally obtained by convolving the noisy signal (Figure B.2c) by the seismograph impulse response (Figure B.5a). The first 0.6 seconds of the noiseless signal filtered by the seismograph (the “recorded” signal) and the noiseless true signal are plotted in Figure B.6. It shows clearly how the true signal is distorted by the seismograph during the recording process.

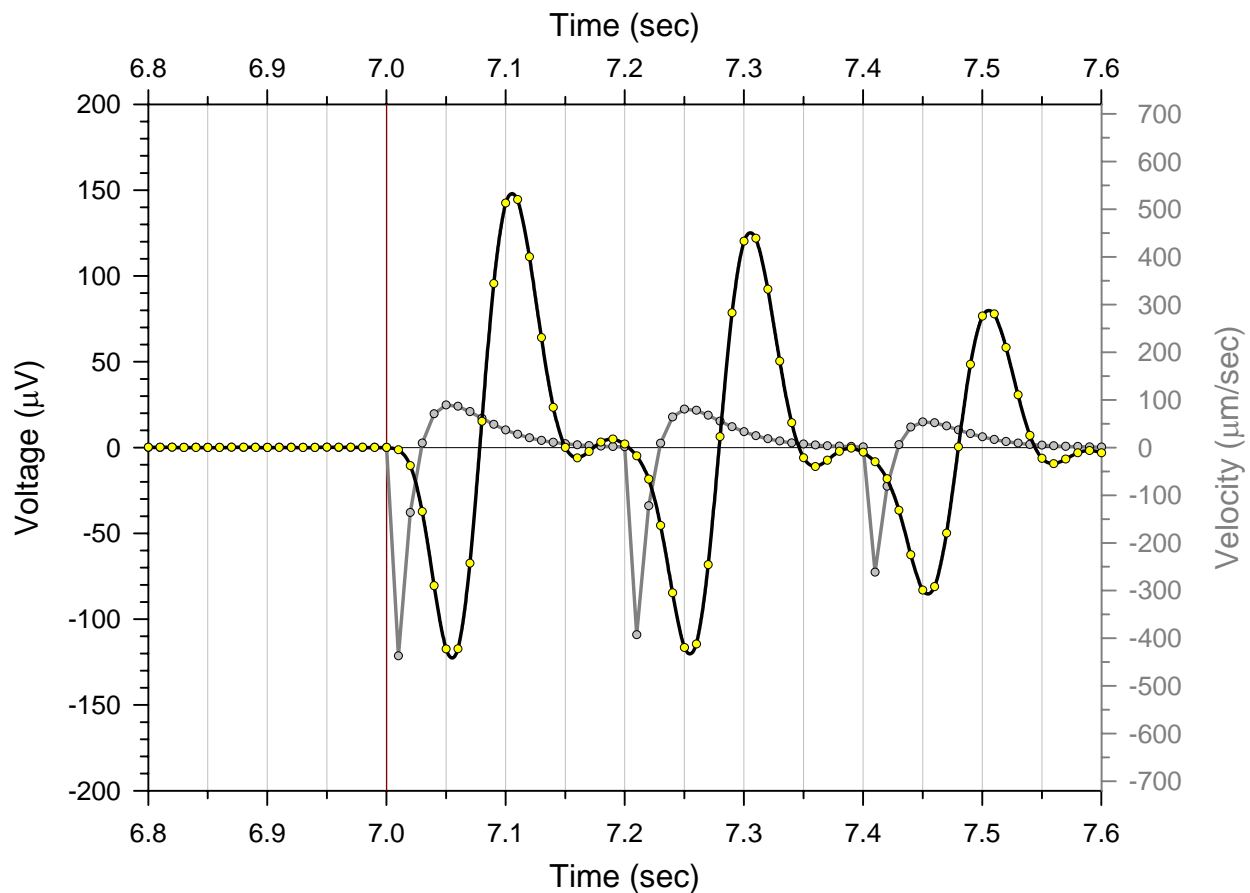


Figure B.6 Noiseless signal (detail). The recorded signal (Voltage) is in black and the corresponding true signal (Velocity) is in gray. The sampling period is 10 msec.

Appendix C

GII and JSO Station Coordinates

C.1 GII STATIONS

GII stations have been primarily located in the past in the (X,Y) Rectangular Israel Grid. Highly accurate formulas from the Survey of Israel were used to convert between the Israel Grid coordinates and Geographical coordinates (Cassini-Soldner projection). Round-trip conversions on reference points are accurate at ± 0.02 m. For field locations where the station name has changed once or several times over the years, only the most recent known name is reported here.

Table C.1 GII station coordinates

Station	X (km)	Y (km)	Elevation (m)	Latitude (°N)	Longitude (°E)
ADI	171.570	276.080	469	33.0797	35.2262
ARAD	178.800	70.000	160	31.2212	35.3018
ARVA	197.550	135.000	-330	31.8071	35.5004
ARVI	168.000	5.600	0	30.6403	35.1886
ASHL	116.500	46.500	290	31.0080	34.6492
ASI	159.200	82.100	850	31.3303	35.0960
ATAR	97.300	48.700	200	31.0268	34.4480
ATR	115.600	42.000	350	30.9674	34.6401
ATZ	175.510	247.590	515	32.8228	35.2682
BGIO	158.490	125.560	760	31.7222	35.0880
BLVR	199.000	222.000	275	32.5917	35.5183
BRNI	149.600	238.100	400	32.7370	34.9918
BSO	101.600	75.550	140	31.2692	34.4912
CRI	153.860	231.990	431	32.6820	35.0373
DAM3	192.350	54.500	-390	31.0812	35.4437
DLIA	156.000	221.000	160	32.5829	35.0603
DOR	120.720	102.120	182	31.5098	34.6907

Table C.1 - continued

Station	X (km)	Y (km)	Elevation (m)	Latitude (°N)	Longitude (°E)
DRGI	187.380	111.060	10	31.5914	35.3925
DS10	187.640	112.020	20	31.6001	35.3953
DS11	197.480	134.460	-373	31.8022	35.4996
DS12	191.930	125.400	9	31.7207	35.4408
DS13	190.490	119.900	169	31.6711	35.4255
DS6	187.770	71.090	-390	31.2309	35.3960
DSD2	191.800	65.000	-390	31.1759	35.4381
DSI	186.600	108.580	200	31.5690	35.3843
EIL	145.000	-101.800	200	29.6712	34.9512
ENGI	187.400	97.600	-300	31.4700	35.3925
GLH	211.230	235.530	335	32.7133	35.6491
GVI	141.510	115.760	389	31.6336	34.9091
GVMR	187.330	219.030	95	32.5651	35.3939
HAF	152.700	242.600	186	32.7776	35.0247
HAMT	147.450	-115.400	0	29.5485	34.9768
HLZ	116.000	112.000	100	31.5987	34.6404
HMDT	198.730	185.120	125	32.2591	35.5143
HNTI	166.000	277.200	250	33.0898	35.1665
HRI	218.760	296.860	1014	33.2660	35.7327
HRSH	175.740	233.650	416	32.6971	35.2706
JVI	182.460	149.030	686	31.9339	35.3412
KER	101.790	45.310	363	30.9965	34.4953
KMTI	123.100	-54.400	450	30.0981	34.7229
KOMT	175.400	59.000	200	31.1220	35.2661
KRPI	202.800	206.000	-250	32.4473	35.5582
KSDI	211.807	288.643	170	33.1922	35.6577
KSHT	226.050	265.710	700	32.9848	35.8090
KZIT	92.200	36.500	200	30.9164	34.3956
MAMI	163.870	212.290	461	32.5044	35.1442
MBH	141.550	-88.400	842	29.7920	34.9152
MGI	190.300	206.400	400	32.4512	35.4253
MKT	164.430	39.720	517	30.9481	35.1512
MMLI	188.840	204.840	508	32.4371	35.4097
MMR	187.970	266.130	1100	32.9898	35.4016
MNFI	172.500	250.750	350	32.8513	35.2361
MNIT	153.740	207.800	118	32.4639	35.0365
MOI	182.800	127.000	450	31.7352	35.3445
MRN	186.600	268.400	900	33.0103	35.3870
MSDA	177.350	80.450	400	31.3154	35.2867
MSH	129.890	45.280	364	30.9975	34.7895
MZDA	184.450	79.500	-300	31.3068	35.3612
NEOT	186.800	43.490	-390	30.9820	35.3853
NOH	146.130	11.250	680	30.6910	34.9603
NSH	148.000	136.000	300	31.8162	34.9771

Table C.1 - continued

Station	X (km)	Y (km)	Elevation (m)	Latitude (°N)	Longitude (°E)
OFRI	149.083	224.881	100	32.6178	34.9865
PRNI	150.300	-26.360	412	30.3518	35.0046
RAMI	161.420	243.500	37	32.7859	35.1178
RMN	115.460	-8.420	1003	30.5126	34.6413
RMNI	127.000	1.000	880	30.5980	34.7611
RTMM	119.590	51.060	259	31.0493	34.6814
SAGI	118.800	-42.400	560	30.2062	34.6777
SDOM	186.660	54.100	-164	31.0777	35.3840
SHLH	105.900	45.050	310	30.9944	34.5383
SHZF	106.250	34.200	300	30.8965	34.5427
SVTA	114.060	38.320	370	30.9341	34.6241
SZAF	107.280	34.510	293	30.8994	34.5534
YAIR	182.000	70.400	80	31.2247	35.3354
YASH	138.200	-112.700	680	29.5727	34.8813
YTIR	160.980	84.620	905	31.3530	35.1146
ZEFH	140.690	-118.320	118	29.5220	34.9072
ZELM	105.500	56.500	215	31.0976	34.5334
ZFRI	166.950	-2.850	0	30.5641	35.1777
ZNT	152.780	182.890	230	32.2392	35.0267

C.2 JSO STATIONS

JSO station coordinates are given by JSO in Geographical coordinates. Conversion to Rectangular Israel Grid coordinates was made with the same conversion formulas as those in C.1.

Table C.2 JSO station coordinates

Station	X (km)	Y (km)	Elevation (m)	Latitude (°N)	Longitude (°E)
AQBJ	154.570	-95.573	170	29.7275	35.0500
DHLJ	188.424	25.538	-80	30.8200	35.4020
HSHJ	187.450	-129.466	1170	29.4217	35.3893
KFNJ	214.156	141.143	-90	31.8620	35.6760
LISJ	195.871	72.117	-327	31.2400	35.4810
MASJ	218.105	126.391	783	31.7288	35.7170
MKRJ	210.979	106.757	815	31.5520	35.6410
MRSJ	180.891	-100.290	810	29.6850	35.3220
NAQJ	198.323	-65.543	1640	29.9982	35.5030
SALJ	214.776	157.469	780	32.0092	35.6833

References

- Akaike, H. (1969) Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21, 243-247.
- Akaike, H. (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, AC-19, 716-723. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis*: IEEE Press, New York, NY, 234-241.
- Aki, K., and Richards, P. (1980) *Quantitative Seismology, Theory and Methods*. W.H. Freeman, San Francisco, CA.
- Aldersons, F., Ben-Avraham, Z., Hofstetter, A., Kissling, E., and Al-Yazjeen, T. (2003) Lower-crustal strength under the Dead Sea basin from local earthquake data and rheological modeling. *Earth and Planetary Science Letters*, 214, 129-142.
- Allen, R. (1978) Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, 68, 1521-1532.
- Allen, R. (1982) Automatic phase pickers: their present use and future prospects. *Bulletin of the Seismological Society of America*, 72, S225-S242.
- Al-Zoubi, A., and Ben-Avraham, Z. (2002) Structure of the earth's crust in Jordan from potential field data. *Tectonophysics*, 346, 45-59.
- Al-Zoubi, A., and ten Brink, U. (2002) Lower crustal flow and the role of shear in basin subsidence: an example from the Dea Sea basin. *Earth and Planetary Science Letters*, 199, 67-79.
- Andersen, N. (1974) On the calculation of filter coefficients for maximum entropy spectral analysis. *Geophysics*, 39, 1, 69-72. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis*: IEEE Press, New York, NY, 252-255.
- Anderson, E.M. (1951) *The Dynamics of Faulting*, 2nd edition. Oliver and Boyd, Edinburgh, England.

- Aster, R., and Rowe, C. (2000) Automatic phase pick refinement and similar event association in large seismic datasets. In Thurber, C.H., and Rabinowitz, N., (Eds.) *Advances in seismic event location: Kluwer Academic Publishers, Dordrecht, Modern approaches in geophysics*, 18, 231-263.
- Baer, M., and Kradolfer, U. (1987) An automatic phase picker for local and teleseismic events. *Bulletin of the Seismological Society of America*, 77, 1437-1445.
- Baggeroer, A.B. (1976) Confidence Intervals for Regression (MEM) Spectral Estimates. *IEEE Transactions on Information Theory*, IT-22, 534-545. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis: IEEE Press, New York, NY*, 150-161.
- Banda, E., and Cloetingh, S. (1992) Physical properties of the Europe's lithosphere. In Blundell, D., Freeman, R., and Mueller, S. (Eds.) *A continent revealed: The European Geotraverse: Cambridge University Press, Cambridge*, 71-80.
- Bartov, Y., Steinitz, G., Eyal, M., and Eyal, Y. (1980) Sinistral movement along the Gulf of Aqaba – its age and relation to the opening of the Red Sea. *Nature*, 285, 220-221.
- Basseville, M., and Nikiforov, I.V. (1993) *Detection of Abrupt Changes: Theory and Application. Prentice Hall Information and System Science Series, Prentice Hall, Englewood Cliffs, NJ.*
- Beardmore, G.R., and Cull, J.P. (2001) *Crustal Heat Flow: a guide to measurement and modeling. Cambridge University Press, Cambridge.*
- Ben-Avraham, Z. (1997) Geophysical framework of the Dead Sea: structure and tectonics. In Niemi, T.M., Ben-Avraham, Z., and Gat, J.R. (Eds.) *The Dead Sea: The lake and its settings: Oxford University Press, New York, NY*, 36, 22-35.
- Ben-Avraham, Z., and Von Herzen, R.P. (1987) Heat flow and continental breakup: the Gulf of Elat (Aqaba). *Journal of Geophysical Research*, 92, 1407-1416.
- Ben-Avraham, Z., Hänel, R., and Villinger, H. (1978) Heat flow through the Dead Sea rift. *Marine Geology*, 28, 253-269.
- Ben-Menahem, A. (1991) Four Thousand Years of Seismicity Along the Dead Sea Rift. *Journal of Geophysical Research*, 96, B12, 20195-20216.
- Berkhout, A.J., and Zaanen, P.R. (1976) A comparison between Wiener filtering, Kalman filtering, and deterministic least squares estimation. *Geophysical Prospecting*, 24, 141-197.
- Blackman, R.B., and Tukey, J.W. (1958) *The Measurement of Power Spectra, Dover, New York, NY.*
- Bode, H.W., and Shannon, C.E. (1950) A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory. *Proceedings of the Institute of Radio Engineers*, 38, 417-425. Reprinted in Sloane, N.J.A., and Wyner, A.D. (Eds.) (1993) *Claude Elwood Shannon Collected Papers: IEEE Press, New York, NY*, 628-636.

- Brune, J.N. (1970) Tectonic Stress and the Spectra of Seismic Shear Waves from Earthquakes. *Journal of Geophysical Research*, 75, 26, 4997-5009.
- Bryant, A.S., and Jones, L.M. (1992) Anomalous Deep Crustal Earthquakes in the Ventura Basin, Southern California. *Journal of Geophysical Research*, 97, 437-447.
- Bullen, K.E. (1960) Note on Cusps in Seismic Travel-Times. *Geophysical Journal of the Royal Astronomical Society*, 3, 354-359.
- Burg, J.P. (1967) Maximum entropy spectral analysis. Paper presented at the 37th Annual Meeting of the Society of Exploration Geophysicists, Oklahoma City, OK. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis*: IEEE Press, New York, NY, 34-41.
- Burg, J.P. (1968) A new analysis technique for time series data. Paper presented at the NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics, Enschede, Netherlands. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis*: IEEE Press, New York, NY, 42-48.
- Burg, J.P. (1975) Maximum entropy spectral analysis. Ph.D. thesis, Stanford University, pp. 123.
- Byerlee, J.D. (1968) Brittle-ductile transition in rocks. *Journal of Geophysical Research*, 73, 4741-4750.
- Byerlee, J.D. (1978) Friction of rocks. *Pure and Applied Geophysics*, 116, 615-626.
- Camelbeek, T., and Iranga, M.D. (1996) Deep crustal earthquakes and geometry of active faults along the Rukwa trough, East Africa. *Geophysical Journal International*, 124, 612-630.
- Chen, W-P., and Molnar, P. (1983) Focal depths of intracontinental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere. *Journal of Geophysical Research*, 88, B5, 4183-4214.
- Childers, D.G. (Ed.) (1978) *Modern spectrum analysis*. IEEE Press, New York, NY.
- Christensen, N.I., and Mooney, W.D. (1995) Seismic velocity structure and composition of the continental crust: a global view. *Journal of Geophysical Research*, 100, B7, 9761-9788.
- Cloetingh, S., and Ben-Avraham, Z. (Eds.) (2002) *From continental extension to collision: Africa-Europe interaction, the Dead Sea and analogue natural laboratories*. EGU Stephan Mueller Special Publication Series, 2.
- Cloetingh, S., and Burov, E. (1996) Thermomechanical structure of European continental lithosphere: constraints from rheological profiles and EET estimates. *Geophysical Journal International*, 124, 695-723.
- Creager, K.C. (1997) Coral. *Seismological Research Letters*, 68, 269-271.
- Dai, H., and MacBeth, C. (1995) Automatic picking of seismic arrivals in local earthquake data using an artificial neural network. *Geophysical Journal International*, 120, 758-774.
- Dai, H., and MacBeth, C. (1997) The application of back-propagation neural network to automatic

- picking seismic arrivals from single-component recordings. *Journal of Geophysical Research*, 102, 15105-15115.
- Deichmann, N., and Rybach, L. (1989) Earthquakes and temperatures in the lower crust below the northern alpine foreland of Switzerland. In Mereu, R.F., Mueller, S., Fountain, D.M., (Eds.) *Properties and processes of the Earth's Lower Crust: Geophysical Monograph 51*, IUGG, Vol. 6, American Geophysical Union, Washington, DC, 197-213.
- Der, Z.A., and Shumway, R.H. (1999) Phase onset time estimation at regional distances using the CUSUM algorithm. *Physics of the Earth and Planetary Interiors*, 113, 227-246.
- Der, Z.A., McGarvey, M.W., and Shumway, R.H. (2000) Automatic interpretation of regional short period seismic signals using the CUSUM-SA algorithm. *Proceedings of the 22nd DoD/DoE Seismic Research Symposium*. New Orleans, LA.
- DESERT Group, Weber, M., Abu-Ayyash, K., Abueladas, A., Agnon, A., Al-Amoush, H., Babeyko, A., Bartov, Y., Baumann, M., Ben-Avraham, Z., Bock, G., Bribach, J., El-Kelani, R., Förster, A., Förster, H.-J., Frieslander, U., Garfunkel, Z., Grunewald, S., Götze, H.J., Haak, V., Haberland, C., Hassouneh, M., Helwig, S., Hofstetter, A., Jäckel, K.-H., Kesten, D., Kind, R., Maercklin, N., Mechie, J., Mohsen, A., Neubauer, F.M., Oberhänsli, R., Qabbani, I., Ritter, O., Rümpker, G., Rybakov, M., Ryberg, T., Scherbaum, F., Schmidt, J., Schulze, A., Sobolev, S., Stiller, M., Thoss, H., Weckmann, U., and Wylegalla, K. (2004) The crustal structure of the Dead Sea Transform. *Geophysical Journal International*, 156, 655-681.
- Déverchère, J., Petit, C., Gileva, N., Radziminovitch, N., Melnikova, V., and San'kov, V. (2001) Depth distribution of earthquakes in the Baikal rift system and its implications for the rheology of the lithosphere. *Geophysical Journal International*, 146, 714-730.
- Di Stefano, R., Aldersons, F., Kissling, E., Chiarabba, C., and Giardini, D. (2004) Automatic seismic phase picking and consistent observation error assessment: Application to the Italian seismicity. In preparation for submission to *Geophysical Journal International*.
- Dodge, D.A., Beroza, G.C., and Ellsworth, W.L. (1995) Foreshock sequence of the 1992 Landers, California earthquake and its implications for earthquake nucleation. *Journal of Geophysical Research*, 100, 9865-9880.
- Douglas, A. (1997) Bandpass Filtering to Reduce Noise on Seismograms: Is There a Better Way? *Bulletin of the Seismological Society of America*, 87, 4, 770-777.
- Douglas, A., Bowers, D., and Young, J.B. (1997) On the onset of P seismograms. *Geophysical Journal International*, 129, 681-690.
- Durbin, J. (1960) The fitting of time series models. *Review of the International Institute of Statistics*, 28, 3, 233-244.

- Eckstein, Y., and Simmons, G. (1978) Measurements and interpretation of terrestrial heat flow in Israel. *Geothermics*, 6, 117-142.
- El-Isa, Z.H. (1990) Lithospheric structure of the Jordan-Dead Sea transform from earthquake data. *Tectonophysics*, 180, 29-36.
- Erickson, A.J. (1970) The measurement and interpretation of heat flow in the Mediterranean and Black Sea. Ph.D. thesis, Massachusetts Institute of Technology, pp. 433.
- Fischer, R.A. (1936) The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, 7, 179-188.
- Fischer, R.A. (1938) The Statistical Utilization of Multiple Measurements. *Annals of Eugenics*, 8, 376-386.
- Franklin, J.N. (1970) Well-posed stochastic extensions of ill-posed linear problems. *Journal of Mathematical Analysis and Applications*, 31, 682-716.
- Freund, R. (1965) A model of the structural development of Israel and adjacent areas since Upper Cretaceous times. *Geological Magazine*, 102, 189-205.
- Freund, R., Garfunkel, Z., Zak, I., Goldberg, M., Weissbrod, T., and Derin, B. (1970) The shear along the Dead Sea rift. *Philosophical Transactions of the Royal Society of London, Series A*, 267, 105-127.
- Frieslander, U., and Ben-Avraham, Z. (1989) Magnetic field over the Dead Sea and vicinity. *Marine and Petroleum Geology*, 6, 148-160.
- Garfunkel, Z. (1997) The history and formation of the Dead Sea basin. In Niemi, T.M., Ben-Avraham, Z., and Gat, J.R. (Eds.) *The Dead Sea: The lake and its settings*: Oxford University Press, New York, NY, 36, 36-52.
- Garfunkel, Z. (1999) History and paleogeography during the Pan-African orogen to stable platform transition: Reappraisal of the evidence from the Elat area and the northern Arabian-Nubian Shield. *Israel Journal of Earth Sciences*, 48, 135-157.
- Gersch, W., and Sharpe, D.R. (1973) Estimation of power spectra with finite order autoregressive models. *IEEE Transactions on Automatic Control*, AC-18, 367-369.
- Ginzburg, A., and Ben-Avraham, Z. (1997) A seismic refraction study of the north basin of the Dead Sea, Israel. *Geophysical Research Letters*, 24, 2063-2066.
- Ginzburg, A., Makris, J., Fuchs, K., and Prodehl, C. (1981) The structure of the crust and upper mantle in the Dead Sea rift. *Tectonophysics*, 80, 109-119.
- Ginzburg, A., Makris, J., Fuchs, K., Perathoner, B., and Prodehl, C. (1979) Detailed Structure of the Crust and Upper Mantle Along the Jordan-Dead Sea Rift. *Journal of Geophysical Research*, 84, B10, 5605-5612.
- Goetze, C. (1978) The mechanisms of creep in olivine. *Philosophical Transactions of the Royal Society of London, Series A*, 288, 99-119.

- Goldstein, P., Dodge, D., and Firpo, M. (1999) SAC2000: Signal processing and analysis tools for seismologists and engineers, UCRL-JC-135963, Invited contribution to the IASPEI International Handbook of Earthquake and Engineering Seismology.
- GSE/JAPAN/40 (1992) A fully automated method for determining the arrival times of seismic waves and its application to an on-line processing system. Paper tabled in the 34th GSE session in Geneva GSE/RF/62.
- Hansen, F.D., and Carter, N.L. (1982) Creep of selected crustal rocks at 1000 MPa. *Eos Transactions American Geophysical Union*, 63, 437.
- Haslinger, F. (1998) Velocity Structure, Seismicity and Seismotectonics of Northwestern Greece Between the Gulf of Arta and Zakynthos. Ph.D. thesis, ETH, Zürich, pp. 160.
- Havskov, J., and Ottemöller, L. (1999) SeisAn earthquake analysis software, *Seismological Research Letters*, 70, 5, 532-534.
- Hayes, M.H. (1996) Statistical digital signal processing and modeling. John Wiley, New York, NY.
- Jackson, J.A. (2002) Strength of the continental lithosphere: Time to abandon the jelly sandwich ? *GSA Today*, 12, 9, 4-10.
- Jaynes, E.T. (1957a) Information Theory and Statistical Mechanics. *The Physical Review*, 106, 4, 620-630.
- Jaynes, E.T. (1957b) Information Theory and Statistical Mechanics. II. *The Physical Review*, 108, 2, 171-190.
- Johnson, C.E., Lindh. A.G., and Hirshorn, B. (1994) Robust regional phase association. U.S.G.S. Open File Report 94-621.
- Johnson, S.J., and Anderson, N. (1978) On power estimation in maximum entropy spectrum analysis. *Geophysics*, 43, 681-690.
- Jones, R.H. (1974) Identification and Autoregressive Spectrum Estimation. *IEEE Transactions on Automatic Control*, AC-19, 894-897. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis: IEEE Press, New York, NY*, 301-304.
- Jones, R.H. (1976) Autoregression Order Selection. *Geophysics*, 41, 771-773. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis: IEEE Press, New York, NY*, 249-251.
- Jordan Seismological Observatory (JSO) (1993) Earthquakes in Jordan and Adjacent Areas, 25, Amman, Jordan, 1-142.
- Joswig, M., and Schulte-Theis, H. (1993) Master-event correlations of weak local earthquakes by dynamic waveform matching. *Geophysical Journal International*, 113, 562-574.
- Khintchine, A. (1934) Korrelationstheorie der stationären stochastischen Prozesse. *Mathematische Annalen*, 109, 604-615.

- Kirby, S.H. (1983) Rheology of the lithosphere. *Reviews of Geophysics and Space Physics*, 21, 6, 1458-1487.
- Kameyama, M., Yuen, D.A., and Karato, S.-I. (1999) Thermal-mechanical effects of low-temperature plasticity (the Peierls mechanism) on the deformation of a viscoelastic shear zone. *Earth and Planetary Science Letters*, 168, 159-172.
- Kissling, E. (1988) Geotomography with local earthquake data. *Review of Geophysics*, 26, 4, 659-698.
- Kissling, E., Ellsworth, W.L., Eberhart-Phillips, D., and Kradolfer, U. (1994) Initial reference models in local earthquake tomography. *Journal of Geophysical Research*, 99, 19635-19646.
- Klumpen, E., and Joswig, M. (1993) Automated reevaluation of local earthquake data by application of generic polarization patterns for P- and S-onsets. *Computers & Geosciences*, 19, 2, 223-231.
- Kromer, R. (1970) Asymptotic properties of the autoregressive spectral estimator. Ph.D. thesis, Stanford University.
- Kushnir, A., Lapshin, V., Pinsky, V., and Fyen, J. (1990) Statistically optimal event detection using small array data. *Bulletin of the Seismological Society of America*, 80, 6b, 1934-1950.
- Lacoss, R.T. (1971) Data adaptive spectral analysis methods. *Geophysics*, 36, 661-675. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis*: IEEE Press, New York, NY, 134-148.
- Landers, T.E., and Lacoss, R.T. (1977) Some Geophysical Applications of Autoregressive Spectral Estimates. *IEEE Transactions on Geoscience Electronics*, GE-15, 26-32. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis*. IEEE Press, New York, NY, 305-311.
- Langston, C.A. (1979) Structure under Mount Rainier, Washington, inferred from teleseismic body waves. *Journal of Geophysical Research*, 84, 4749-4762.
- Leonard, M. (2000) Comparison of Manual and Automatic Onset Time Picking. *Bulletin of the Seismological Society of America*, 90, 6, 1384-1390.
- Leonard, M., and Kennett, B.L.N. (1999) Multi-component autoregressive techniques for the analysis of seismograms. *Physics of the Earth and Planetary Interiors*, 113, 247-263.
- Levinson, N. (1947a) The Wiener RMS error criterion in filter design and prediction. *Journal of Mathematics and Physics*, 25, 4, 261-278. Reprinted In Wiener, N. (1949) *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*: MIT Press, Cambridge, MA, 129-148.
- Levinson, N. (1947b) A heuristic exposition of Wiener's mathematical theory of prediction and filtering. *Journal of Mathematics and Physics*, 26, 2, 110-119. Reprinted In Wiener, N.

- (1949) Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications: MIT Press, Cambridge, MA, 149-160.
- Mackwell, S.J, Zimmerman, M.E., and Kohlstedt, D.L. (1998) High-temperature deformation of dry diabase with applications to tectonics on Venus. *Journal of Geophysical Research*, 103, B1, 975-984.
- Maggi, A., Jackson, J.A., McKenzie, D., and Priestley, K. (2000) Earthquake focal depths, effective elastic thickness, and the strength of the continental lithosphere. *Geology*, 28, 6, 495-498.
- Marple, S.L.Jr. (1982) Frequency resolution of Fourier and maximum entropy spectral estimates. *Geophysics*, 47, 9, 1303-1307.
- McKenzie, D., and Fairhead, D. (1997) Estimates of the effective elastic thickness of the continental lithosphere from Bouguer and free air gravity anomalies, *Journal of Geophysical Research*, 102, 27523-27552.
- Meissner, R., and Strehlau, L. (1982) Limits of stresses in continental crusts and their relation to the depth-frequency distribution of shallow earthquakes. *Tectonics*, 1, 73-89.
- Morita, Y., and Hamaguchi, H. (1984) Automatic detection of onset time of seismic waves and its confidence interval using the autoregressive model fitting. *Zisin*, 37, 281-293.
- Neev, D., and Hall, J. (1979) Geophysical investigations in the Dead Sea. *Sedimentary Geology*, 23, 209-238.
- Nyblade, A.A., and Langston, C.A. (1995) East African earthquakes below 20 km and their implications for crustal structure. *Geophysical Journal International*, 121, 49-62.
- Oppenheim, A.V., Schaffer, R.W., and Buck, J.R. (1999) *Discrete-time signal processing*, 2nd edition. Prentice Hall, Upper Saddle River, N.J.
- Park, J., Lindberg, C.R., and Vernon III, F.L. (1987) Multitaper Spectral Analysis of High-Frequency Seismograms. *Journal of Geophysical Research*, 92, B12, 12675-12684.
- Pe'eri, S., Wdowinski, S., Shtibelman, A., Bechor, N., Bock, Y., Nikolaidis, R., and van Domselaar, M. (2002) Current plate motion across the Dead Sea Fault from three years of continuous GPS monitoring. *Geophysical Research Letters*, 10.1029/2001GL013879.
- Penland, C., Ghil, M., and Weickmann, K.M. (1991) Adaptive Filtering and Maximum Entropy Spectra With Application to Changes in Atmospheric Angular Momentum. *Journal of Geophysical Research*, 96, D12, 22659-22671.
- Peraldi, R., and Clement, A. (1972) Digital processing of refraction data study of first arrivals. *Geophysical Prospecting*, 20, 529-548.
- Percival, D.B., and Walden, A.T. (1993) *Spectral Analysis for Physical Applications: multitaper and conventional univariate techniques*. Cambridge University press.

- Piersanti, A., Nostro, C., and Riguzzi, F. (2001) Active displacement field in the Suez-Sinai area: the role of postseismic deformation. *Earth and Planetary Science Letters*, 193, 13-23.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1992) *Numerical Recipes in Fortran: The Art of Scientific Computing*, 2nd edition. Cambridge University Press, Cambridge, England.
- Prodehl, C., Fuchs, K., and Mechie, J. (1997) Seismic-refraction studies of the Afro-Arabian rift system – a brief review. *Tectonophysics*, 278, 1-13.
- Quennell, A.M. (1958) The structural and geomorphic evolution of the Dead Sea rift. *Quarterly Journal of the Geological Society of London*, 114, 1-24.
- Rabiner, L.R., and Schafer, R.W. (1978) *Digital Processing of Speech Signals*. Prentice-Hall, Englewood Cliffs, NJ.
- Rabinowitz, N., Steinberg, J., and Mart, Y. (1996) New evidence of magmatic diapirs in the intermediate crust under the Dead Sea, Israel. *Tectonics*, 15, 2, 237-243.
- Rabinowitz, N., and Mart, Y. (2000) Seismic tomography of the Dead Sea region: thinned crust, anomalous velocities and possible magmatic diapirism. In Vendeville, B., Mart, Y., and Vigneresse, J.-L.: *Salt, shale, and Igneous Diapirs in and around Europe*, Geological Society, Special Publications, London, 174, 79-92.
- Ranalli, G. (1987) *Rheology of the Earth: Deformation and flow processes in geophysics and geodynamics*. Allen and Unwin, Boston, MA.
- Ranalli, G., and Murphy, D.C. (1987) Rheological stratification of the lithosphere. *Tectonophysics*, 132, 281-295.
- Rodgers, J.J.W., and Adams, J.A.S. (1969) Uranium and Thorium. In Wedepohl, K.H. (Ed.) *Handbook of Geochemistry*, Vol. II/1: Springer, Berlin, 92K1-92K8 and 90K1-90K3.
- Rowe, C.A., Aster, R.C., Borchers, B., and Young, C.J. (2002) An Automatic, Adaptive Algorithm for Refining Phase Picks in Large Seismic Data Sets. *Bulletin of the Seismological Society of America*, 92, 5, 1660-1674.
- Roy, M., and Royden, L.H. (2000a) Crustal rheology and faulting at strike-slip plate boundaries 1. An analytic model. *Journal of Geophysical Research*, 105, B3, 5583-5597.
- Roy, M., and Royden, L.H. (2000b) Crustal rheology and faulting at strike-slip plate boundaries 2. Effects of lower crustal flow. *Journal of Geophysical Research*, 105, B3, 5599-5613.
- Rybach, L., and Buntebarth, G. (1984) The variation of heat generation, density and seismic velocity with rock type in the continental lithosphere. *Tectonophysics*, 103, 335-344.
- Scherbaum, F. (1996) *Of Poles and Zeros: Fundamentals of Digital Seismometry*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

- Scherbaum, F. and Johnson, J. (1992) Programmable Interactive Toolbox for Seismological Analysis (PITSA). IASPEI Software Library Volume 5, Seismological Society of America, El Cerrito.
- Scholz, C.H. (2002) *The Mechanics of Earthquakes and Faulting*, 2nd edition. Cambridge University Press, Cambridge, England.
- Shannon, C.E. (1948) A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 379-423 (July) and 623-656 (October). Reprinted in Sloane, N.J.A., and Wyner, A.D. (Eds.) (1993) *Claude Elwood Shannon: collected papers*: IEEE Press, New York, NY, 5-83.
- Shapira, A. (1997) On the seismicity of the Dead Sea basin. In Niemi, T.M., Ben -Avraham, Z., and Gat, J.R. (Eds.) *The Dead Sea: The lake and its settings*: Oxford University Press, New York, NY, 36, 82-86.
- Shapira, A., Avni, R., and Nur, A. (1993) A note: New estimate of the Jericho earthquake epicenter of July 11, 1927. *Israel Journal of Earth Sciences*, 42, 93-96.
- Shearer, P.M. (1997) Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence. *Journal of Geophysical Research*, 102, 8269-8283.
- Shelton, G., and Tullis, J.A. (1981) Experimental flow laws for crustal rocks (abstract). *Eos Transactions American Geophysical Union*, 62, 396.
- Shudofsky, G., Cloetingh, S., Stein, S., and Wortel, R. (1987) Unusually deep earthquakes in East Africa: Constraints on the thermo-mechanical structure of a continental rift system. *Geophysical Research Letters*, 14, 741-744.
- Schuster, A. (1898) On the investigation of hidden periodicities with application to a supposed twenty-six day period of meteorological phenomena. *Terrestrial Magnetism*, 3, 1, 13-41.
- Sibson, R.H. (1974) Frictional constraints on thrust, wrench and normal faults. *Nature*, 249, 542-544.
- Sleeman, R., and van Eck, T. (1999) Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings. *Physics of the Earth and Planetary Interiors*, 113, 265-275.
- Smylie, D.E., Clarke, G.K.C., and Ulrych, T.J. (1973) Analysis of irregularities in the earth's rotation. In Bolt, B.A. (Ed.) *Methods in computational physics*, vol. 13: Academic Press, New York, NY, 391-430.
- Stiller, M., Yanaki, N.E., and Kronfeld, J. (1985) Comparative isotope study of two short sediment cores from the Dead Sea. *Chemical Geology (Isotope Geoscience Section)*, 58, 107-119.

- Takanami, T., and Kitagawa, G. (1988) A new efficient procedure for the estimation of onset times of seismic waves. *Journal of Physics of the Earth*, 36, 267-290.
- Takanami, T., and Kitagawa, G. (Eds.) (2003) *Methods and Applications of Signal Processing in Seismic Network Operations*. Springer, Berlin, Lecture Notes in Earth Sciences, 98.
- ten Brink, U.S. (2002) Corrigendum to 'Lower crustal flow and the role of shear in basin subsidence: an example from the Dead Sea basin'. *Earth and Planetary Science Letters*, 201, 447-448.
- ten Brink, U.S., and Ben-Avraham, Z. (1989) The anatomy of a pull-apart basin: seismic reflection observations of the Dead Sea basin. *Tectonics*, 2, 333-350.
- ten Brink, U.S., Kovach, R.L., Schoenberg, N., and Ben-Avraham, Z. (1990) Uplift and a possible Moho offset across the Dead Sea transform. *Tectonophysics*, 180, 71-85.
- ten Brink, U.S., Ben-Avraham, Z., Bell, R., Hassouneh, M., Coleman, D., Andreasen, G., Tibor, G., and Coakley, B. (1993) Structure of the Dead Sea pull-apart basin from gravity analysis. *Journal of Geophysical Research*, 98, 21877-21894.
- Thomson, J.D. (1982) Spectrum estimation and harmonic analysis. *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE)*, 70, 1055-1096.
- Thurber, C.H. (1984) SIMUL3. In Engdahl, E.R. (Ed.) *Documentation of Earthquake Algorithms: Report SE-35*, World Data Center A for Solid Earth Geophysics, Boulder, Colorado, 15-17.
- Turcotte, D.L., and Schubert, G. (2002) *Geodynamics*, 2nd edition. Cambridge University Press, Cambridge, England.
- Turcotte, T., and Arieh, E. (1988) Catalog of earthquakes in and around Israel. In *Preliminary safety analysis report: Israel Electric Corporation*, Tel Aviv, 1-18.
- Ulrych, T.J., and Bishop, T.N. (1975) Maximum Entropy Spectral Analysis and Autoregressive Decomposition. *Revue of Geophysics and Space Physics*, 13, 183-200. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis: IEEE Press*, New York, NY, 54-71.
- van den Bos, A. (1971) Alternative Interpretation of Maximum Entropy Spectral Analysis. *IEEE Transactions on Information Theory*, IT-17, 493-494. Reprinted in Childers, D.G. (Ed.) (1978) *Modern spectrum analysis: IEEE Press*, New York, NY, 92-93.
- van Eck, T., and Hofstetter, A. (1989) Microearthquake activity in the Dead Sea region. *Geophysical Journal International*, 99, 605-620.
- van Eck, T., and Hofstetter, A. (1990) Fault geometry and spatial clustering of microearthquakes along the Dead Sea-Jordan rift fault zone. *Tectonophysics*, 180, 15-27.
- Vaseghi, S. (2000) *Advanced Digital Signal Processing and Noise Reduction*, 2nd edition. John Wiley, Chichester, West Sussex, England.
- Walker, G. (1931) On periodicity in series of related terms. *Proceedings of the Royal Society of*

- London, Series A, 131, 518-532.
- Wiener, N. (1930) Generalized harmonic analysis. *Acta Mathematica*, 55, 117-258.
- Wiener, N. (1949) *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. MIT Press, Cambridge, MA.
- Wong, I., and Chapman, D. (1990) Deep intraplate earthquakes in the western United States and their relationship to lithospheric temperatures. *Bulletin of the Seismological Society of America*, 80, 589-599.
- Wu, N. (1997) *The maximum entropy method*. Springer, Berlin.
- Yule, G.U. (1927) On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London, Series A*, 226, 267-298.

Acknowledgments

I want to particularly thank my two advisors Prof. Zvi Ben-Avraham from Tel Aviv University and Prof. Edi Kissling from ETH Zürich. I knew from the start that I would be in good hands working with them, and I can truly say that this feeling only grew over the years. As I moved my non-portable computer to three continents before concluding the work, I am also very grateful to them for their trust. Since most of my time was spent working from home, I really enjoyed and I miss now our Monday afternoon sessions “on the roof” at TAU where every week members of the Minerva Dead Sea Research Center share the progress of their work with the others. I want to thank Michael Lazar for his translations to or from Hebrew and for having paid my tuition fees when I was living far from the Middle East. Thanks a lot also to Uri Schattner for his meticulous translation to Hebrew of the cover page and abstract of the thesis.

I particularly thank also the Ministry of Energy and Infrastructure of Israel and Dr. Yossi Bartov for the financial support of this work during two years.

I got plenty of seismological data from the Geophysical Institute of Israel, and I thank Dr. Avi Shapira for this opportunity. I particularly thank Dr. Abraham Hofstetter for his interest, involvement and help in this work. I thank also Dr. Nitzan Rabinowitz for having kindly allowed me to attend the Shalheveth Freier First International Workshop on Advanced Methods in Seismic Analysis, held at the Dead Sea in January 1998. Alona Malitzky introduced me to the binary format of seismograms at GII. I greatly thank Lea Feldman and Batia Reich for their help to retrieve, and for retrieving so many of the seismograms. Corinne Ben-Sasson helped to solve many practical issues.

I got also plenty of data from the Jordanian Seismological Observatory for which I am very thankful to Dr. Abdel-Qader Amrat. I particularly thank Dr. Tawfiq Al-Yazjeen for his interest, commitment and help along the road of the project. The few trips I did to NRA in Amman were always fruitful thanks to him. What a surprise also to meet Bassam Al-Biss again, after we briefly worked together in 1988 in the oil exploration of the Djebel Druze region in Jordan.

Discussions with Dr. Manfred Baer were very useful to better understand how his picking algorithm works and I thank him also for providing his valuable code as a freeware contribution to the seismological community. I thank Dr. E.R. Engdahl, Dr. W.L. Ellsworth and Dr. W. Kohler for having provided me with the source code of the R. Allen picking system.

Raffaele di Stefano endured with good will a lot of beta testing of MannekenPix before version 1.6.2 allowed him to work in production mode on his data sets of Italy. Thanks for your patience Raffaele.

Most of the library work for the lower-crustal study of the Dead Sea was done at Lamont-Doherty. I thank Dr. Michael Steckler for granting me access to the library and Mary Ann Brueckner for her very kind help on the spot.

I thank also André Van Moer who introduced me to Matlab and to the Watcom compilers 10 years ago in central Africa. I still use them today. Frances and Russ improved the Lower-crustal manuscript. Our nanny Jessica and “Nonna” Iris did a pretty number of extra hours to take care of the kids.

But of course Marilena’s support was essential. Without my spouse’s natural resistance from earthquake-prone Friuli and financial support when no funding was available, this work could never have been completed. For Elisa and Francesco, the most important thing is probably that dad is back home now.

I finally thank Prof. Sierd Cloetingh, Prof. George Poupinet, and an anonymous reviewer for having accepted to review the thesis and for their constructive criticism.

Thanks to all of you.