Chapter 4

A new absolute arrival time data set for Europe

The main aim of this study is to create a data set of accurate absolute arrival times for stations in Europe which do not report to the ISC (International Seismological Centre). Waveforms were obtained from data centers and temporary experiments and an automatic picking method was applied to determine absolute arrival times for first and later arriving P and S phases. 86,600 arrival times were picked whose distribution of residuals shows generally low standard deviations. Furthermore, mean teleseismic station residuals reflect the properties of the underlying crust and uppermost mantle. Comparison to ISC data for matching event-station-phase combinations also confirms the good quality of the new absolute arrival time picks. Most importantly, this data set complements the ISC data as it fills regional data gaps in Europe.

4.1 Introduction

Arrival times are routinely reported by many seismological networks to the ISC, resulting in bulletins of millions of arrival times since 1964. Clearly, a wealth of information can be gained from these data regarding the Earth's interior, for example, by application of travel time tomography. However, the reporting stations are not distributed equally over the globe therefore leaving gaps, in particular, in the oceans and stable cratonic regions. Furthermore, the quality of these data, which are mostly handpicked, varies greatly (Röhm *et al.*, 1999).

Besides stations included in arrival time bulletins, a large number of seismic stations exist whose waveforms are not used routinely but are sent to data centers for digital storage. For many events included in these waveforms, arrival times were either not picked at all or only with limitations (e.g. a restricted period in time or limited epicentral distance range).

Another valuable source of data is provided by temporary experiments. To fill the geographical gaps, many regional experiments were carried out during the last 15 to 20 years where spatially dense temporary networks were placed in the field for several months. Often, arrival times for events registered at those arrays were only picked relatively. That means, not the

arrival time of a phase onset was determined but the arrival time of the first maximum or minimum after the onset. This procedure has the advantage that observational errors due to high noise levels can be reduced but as a major disadvantage, arrival times are only obtained with respect to the unknown mean network arrival time for a specific event. Therefore, they cannot be used for event relocation or to obtain absolute velocity information on the crust and mantle below the array. Consequently, obtaining arrival times for events recorded at such stations which do not report arrival times to the ISC can provide new detailed information for high-resolution travel time tomography.
Besides using additional stations, the picks should also be of a consistent good quality as erroneous picks will affect or overprint velocity structures which would otherwise be imaged in travel time tomography. Generally, hand picks are considered to give the best quality but for large data sets this approach is not feasible. However, advanced automatic picking techniques can be applied when many waveforms are recorded for a single event which provide very accurate picks.
The aim of this study is to present such a data set of absolute arrival times for Europe. This data set provides high quality picks of previously unused waveforms and can, for instance, be combined with the ISC arrival times for travel time tomography. To ensure a good picking quality for a large number of waveforms we apply a recently developed two-step automatic phase picker. With our approach, we focus on Europe where many temporary experiments have taken place and where data centers provide a large source of additional waveforms from various digital networks in Europe.

4.2 Data

During the past 15 years large teleseismic experiments were conducted in Europe where spatially dense seismograph arrays were placed in the field for several months. Data from such experiments were obtained as waveforms from the CALIXTO, EIFEL, MIDSEA, SVEKALAPKO, and TOR experiment. Another data collection (named ”Leeds” data set hereafter) was provided by the University of Leeds, UK (Arrowsmith, 2003) and the ORFEUS (Observatories and Research Facilities for EUropean Seismology) data center forms a further important source of waveforms. A map of all station locations is displayed in Figure 4.1.

4.2.1 Leeds Data Set

This waveform collection comprises 150 stations from the BGS (British Geological Survey), DIAS (Dublin Institute for Advanced Studies) and LDG (Laboratoire de Détection et de Géophysiqe, France) seismic networks. Many of these stations did not report arrival times to the ISC on a regular basis. The data set contains registrations from the period 1993-2001. It was used so far only for classic teleseismic travel time tomography with relative arrival time picks (Arrowsmith, 2003; Arrowsmith *et al.*, 2005) to investigate the relation between asthenosphere, lithosphere and crust beneath the British Isles. We used the waveforms as provided to us by S. Arrowsmith (pers. comm.).

4.2.2 CALIXTO

The CALIXTO (Carpathian Arc LIthosphere X-TOmography) experiment was carried out to investigate the lithosphere/asthenosphere structure of the Vrancea zone (southeast Carpathians) known for its strong and localized seismicity. From May 1999 to November 1999, 110 mobile stations were placed in the field in Romania. The data set was supplemented by registrations from 18 permanent stations in Romania (see Wenzel *et al.* (1998) for details). Among the published tomography studies about the CALIXTO experiment are teleseismic P tomography studies using relative travel time residuals (Martin and Ritter, 2005; Martin *et al.*, 2006), an upper crustal absolute P tomography (Landes *et al.*, 2004) and a study using handpicked absolute P arrival times by Weidle *et al.* (2005). However, absolute arrival times were neither picked for S waves nor for the entire set of local and regional events.

4.2.3 EIFEL

The EIFEL project was conducted to investigate the Quarternary volcanism in the Eifel and a possible mantle plume as its origin. Between November 1997 and June 1998, 158 stations were operated in the Eifel and surrounding regions. Ritter *et al.* (2000) give an overview of the experiment. Among studies about this experiment are a receiver function study (Grunewald *et al.*, 2001), a teleseismic P tomography (Ritter *et al.*, 2001) and a teleseismic S tomography study (Keyser *et al.*, 2002) both using relative travel time residuals.

4.2.4 MIDSEA

The MIDSEA (Mantle Investigation of the Deep Suture between Eurasia and Africa) project was performed to fill gaps in the Mediterranean area where no broadband registrations existed before and should therefore improve images of the lithosphere and mantle beneath the Mediterranean region (van der Lee *et al.*, 2001). Registrations from 10 stations were available via ORFEUS. These stations had been placed in the field for 1–2 years during the period June 1999 – May 2002. Among studies about this project are a receiver function study (van der Meijde *et al.*, 2003), a surface wave tomography study (Marone *et al.*, 2003) and a shear wave splitting analysis (Schmid *et al.*, 2004). For this data set, absolute arrival times were not picked.

4.2.5 SVEKALAPKO

The SVEKALAPKO (SVEcofennian-KArelian-LAPland-KOla) project was carried out in Finland with the aim to get a better understanding of the formation of the oldest continents, namely the core of the Karelian province of Archean age (2.6 Ga). 128 mobile stations were placed in the field from August 1998 to May 1999. The data set was completed by registrations from 15 permanent stations (see Bock and the SVEKALAPKO Seismic Tomography Working Group (2001) for experiment details). Among studies about this experiment are a receiver function study (Alinaghi *et al.*, 2003), a teleseismic P tomography study (Sandoval *et al.*, 2003, 2004b), a surface wave tomography study (Bruneton *et al.*, 2004) and a local tomography study (Yliniemi *et al.*, 2004). Only for the local study, absolute arrival times were used.

4.2.6 TOR

For the TOR (Teleseismic Tomography across the Tornquist Zone in Germany–Denmark–Sweden) experiment, 120 stations were placed in South Sweden, Denmark and North Germany from October 1996 to April 1997 to image the Tornquist zone, that separates the Baltic shield from the younger (Phanerozoic) parts of Central Europe, in greater detail than before. Many studies already exist for this experiment using a range of seismological methods. Besides an overview of the experiment given by Gregersen *et al.* (2002), among the studies are a teleseismic P tomography (Arlitt *et al.*, 1999), non-linear P and S tomography studies (Shomali *et al.*, 2006; Voss *et al.*, 2006), receiver function studies (Gossler *et al.*, 1999; Wilde-Piórko *et al.*, 2002; Alinaghi *et al.*, 2003) and anisotropy analyses (Wylegalla and TOR Working Group, 1999; Plomerova and TOR Working Group, 2002). However, none of the named studies used absolute travel times.

4.2.7 ORFEUS

The ORFEUS Data Center provides the biggest data set particularly picked for our study. Their archive contains registrations from the years 1988 to 2000 for stations of the European digital seismometer network of which approximately half of the stations did either not report at all or at least not regularly to the ISC.

4.3 Method

The data described above were obtained from the individual sources as waveform registrations. Since the main interest of this study are the arrival times, those waveforms were processed further. They were either obtained already sorted by events or if necessary the events were selected from the EHB catalog (Engdahl *et al.*, 1998), a reprocessed version of the ISC bulletins, using within Europe earthquakes with $m_b > 4.5$ and for teleseismic events earthquakes with $m_b > 5.5$. As the seismograms were provided in various data formats, they were converted by us to the common format SAC (Goldstein *et al.*, 2003). Since the stations were equipped with different types of sensors, which are sensitive in different frequency ranges, the registrations were restituted to simulate the short period WWSSN (World Wide Standardized Seismographic Network) sensor with a dominant frequency around 1 Hz. Besides P-wave arrivals for events in all distance ranges, also S-wave arrivals were picked for the CALIXTO data set. In those cases, the registrations were restituted with a Wiechert sensor with a dominant frequency of 0.1 Hz and continuing high amplification towards higher frequencies to account for the lower frequencies of S-waves. As a next step, the waveforms were bandpass filtered according to the epicentral distance of the registered phase and the phase type (cf. Sandoval *et al.*, 2004a). Theoretical travel times were computed for all registrations in the Earth reference model ak135 (Kennett *et al.*, 1995) to choose appropriate time windows for the travel time picking.

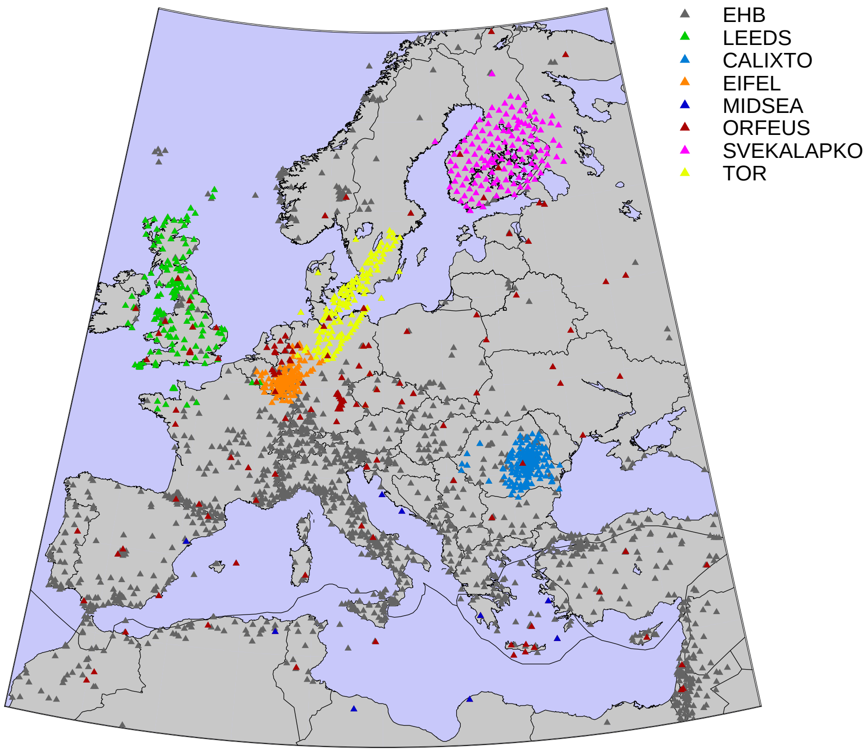

Figure 4.1: Station locations. The different colours denote to which network/data source the stations belong.

Because of the large number of waveforms, hand-picking was not feasible anymore. Therefore, the picking of the arrival times was carried out with a semi-automated picking software developed by Sandoval *et al.* (2004a). The picking was performed per event with a two-step algorithm to ensure a good picking quality. First, the STA/LTA algorithm of Earle and Shearer (1994) was applied. This algorithm is based on a short-term-average to long-term-average ratio taken along an envelope function of the seismogram and returns the absolute arrival time as a result. Then, the picks were grouped regionally and a reference station with a high signal-to-noise ratio was selected manually for each region. The reference waveforms were cross-correlated with each other for consistency between the regions. Afterwards, the

waveforms within each region were cross-correlated with the reference station to improve the STA/LTA pick. This method was applied for picking of the CALIXTO, EIFEL, SVEKALAPKO and TOR experiments.
Later on, the procedure was refined as it required input from the user to choose the reference waveforms which would have been too time-consuming for larger data sets. Additionally, as grouping by equally sized subregions is not the optimum approach for wide-spread station locations as the ORFEUS stations, a cluster analysis was performed to group the stations. Instead of cross-correlation an adaptive stacking method based on the algorithm of Rawlinson and Kennett (2004) was applied to each cluster as this method is more robust concerning waveform variability than cross-correlation techniques. With this method, all waveforms were initially aligned with respect to their theoretical travel time and stacked. They were shifted with respect to the stack to iteratively maximize the stack. As a result, the travel time residual of each phase relative to its theoretical travel time was obtained and an error estimate was computed via the misfit between the individual waveforms and the stacked signal. Subsequently, the STA/LTA picker was applied to the stacked waveform to obtain an accurate pick of the phase onset and finally absolute arrival times were computed for each station from the combination of STA/LTA pick and relative station residuals. The refined method was applied to the ORFEUS, MIDSEA and Leeds data sets.
Such methods work well if the waveforms within a group/cluster are similar to each other but work less well if waveforms change rapidly with distance as can be the case for example for local events. Then, the picker allows for user input to pick the arrival times by hand.

4.4 Results

In total, 86580 absolute arrival time picks were obtained with an estimated picking error of 0.10–0.20 s. Table 4.1 gives a more detailed overview of the number of picks. All residuals used and displayed in this study are computed with respect to ak135. As is shown in Figure 4.2, the residuals (observed – theoretical travel time) are centered around -0.13 s with a standard deviation of 1.05 s appropriate for a data set that contains information about many different geological settings. The solid black line in Figure 4.2 represents the best fitting Gaussian curve with a standard deviation of 0.67 s. Towards bigger residuals, the observed data do not follow a normal distribution but show a broader tail. This effect is mainly attributed to errors in the data set (for example event mislocation or picking errors) but also to 3-D velocity structures along the ray path (e.g. Röhm, 1999; Pulliam *et al.*, 1993).

Generally, rays from teleseismic events recorded at a dense station array take approximately the same path through the Earth except for the part directly beneath the array. Therefore, their travel time differences for each event reflect the velocity differences of the crust and lithosphere beneath the array. Two examples which indicate the high quality of the obtained picks are illustrated in Figure 4.3. On the left, an event in the Afghanistan-Tajikistan border region registered at the EIFEL array is displayed. It shows highest residuals in the center of the array and west of it (as the wavefront arrives from the east) indicating the lower velocities directly beneath the Eifel. On the right, the obtained travel time residuals for an event in Japan registered at the TOR array are shown. A transition can be observed from negative

Data Set	Picks	Events	Distance Range
Leeds	5237 (P)	64	28°– 162°
CALIXTO	4078 (P)	210	0°– 158°
	1484 (SV)	100	0°– 86°
	1564 (SH)	101	0°– 88°
EIFEL	6288 (P)	90	5°– 164°
MIDSEA	739 (P)	256	0°– 160°
ORFEUS	56686 (P)	2056	0°– 178°
SVEKALAPKO	5606 (P)	102	20°– 148°
TOR	4898 (P)	111	2°– 160°

Table 4.1: Summary of the obtained arrival time picks.

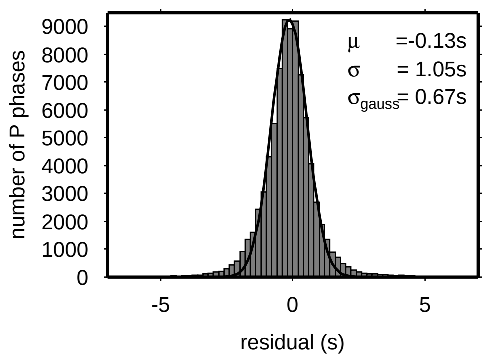

Figure 4.2: Histogram of the P residuals (with respect to ak135) for the newly picked data. The solid line indicates the best fitting Gaussian function.

residuals in the north due to the high velocities of the Baltic shield across the Tornquist zone to higher residuals related to the slower velocities of the younger parts of central Europe.

Since the automatic picker works best for teleseismic events, even though data for all distance ranges were picked, for the remainder of this paper, focus will be on arrival times obtained for teleseismic events ($\Delta > 28°$). Computing the mean of all teleseismic residuals obtained at a single station will mainly reflect the regional velocity variations underneath the stations as for a wide azimuthal coverage of epicentral regions source effects and contributions of the paths further away from the station will diminish. In Figure 4.4, the mean teleseismic station residuals of the new data and their standard deviations are presented (upper and lower right respectively). For comparison, the mean teleseismic residuals and standard deviations of the ISC bulletins are displayed on the upper and lower left.

Negative residuals are found beneath the SVEKALAPKO array and the northern part of the TOR array reflecting the high velocities of the Baltic shield. The residuals become positive at the southern part of the TOR array due to the lower velocities of the underlying lithosphere.

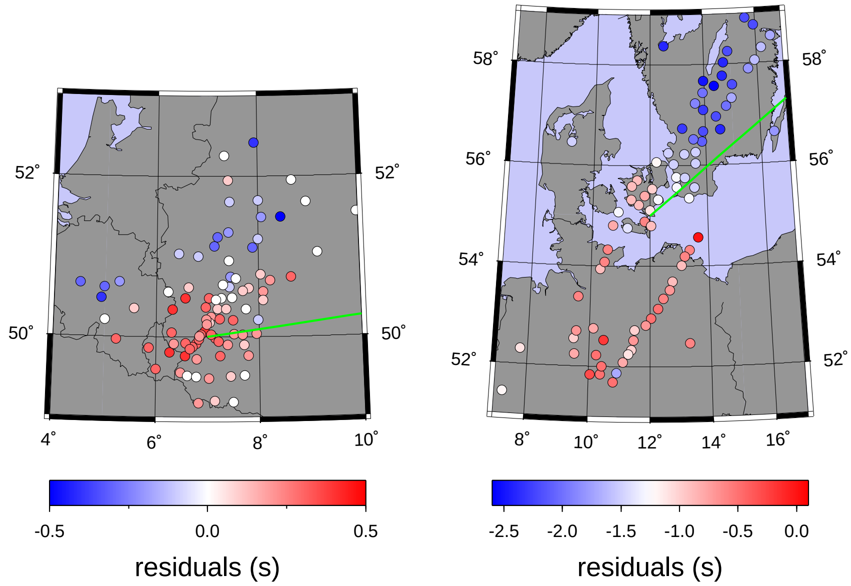

Figure 4.3: Example of residuals (with respect to ak135) obtained for the EIFEL (left) and TOR array (right) for an event in the Afghanistan-Tajikistan border region (30 May 1998 06:22:31.1, 37.141°N, 70.094°E, z = 35.2 km, m_b = 5.7) and Japan (19 October 1996 14:44:41.75, 31.911°N, 131.574°E, z = 18 km, m_b = 6.2) respectively. The green line indicates the direction of incidence of the wavefront.

The Eifel region shows only in the central part slightly positive residuals (on average) related with the lower velocities underneath it but otherwise negative velocities due to higher velocities in the surrounding crust and lithosphere (particularly to the north and northeast). For the British data set, generally small residuals are found with the most negative residuals in the Central Highlands region possibly caused by high velocities of remnants of a subducted oceanic plate (Arrowsmith, 2003, and reference therein). The CALIXTO array shows positive residuals in the bend zone of the southeast Carpathians and in the Transylvanian and Focsani basins. Negative residuals are found towards the northeast on the East European Platform, the east and in the southwestern part of the seismic array (west of the Intra-Moesian fault). Overall, 50 of the 160 stations in the ORFEUS archive did not report arrival times to the ISC, in particular NARS stations (Network of Autonomously Recording Seismographs, see e.g. Paulssen *et al.* (1990, 2000)) and about 40 more stations either reported only local and regional arrival times to the ISC or not for the entire period of operation. Picking of arrival times from the rest of the stations did not consume much extra time due to the applied picking method. The ORFEUS stations on the East European Platform display negative residuals due to the fast velocities of the old cratonic material beneath it while for example the stations in the Netherlands show positive residuals.

A comparison of the new data with the mean teleseismic residuals of the EHB catalog (orig-

inating from the ISC bulletins) shows similarities but due to the locally denser station distribution regional variations can be seen in more detail as for example across the Tornquist zone.
The standard deviation of the new teleseismic residuals are generally low around 0.5–0.7 s except for regions where the residuals show a strong azimuthal dependence (e.g. TOR - many events either from north-northeast along array direction or perpendicular to the array from east-southeast, CALIXTO - complex tectonic structure due to collisional setting, deep sediment basins) or concerning the highest standard deviations where only few teleseismic arrival time picks with greater variation exist. Nevertheless, the scatter of residuals is much lower than for most of the ISC data indicating the high quality of the data set.
As a final inspection of the new picks, identical teleseismic event-station-phase pairs were retrieved from the EHB catalog and compared to the new picks. Approximately 17,000 such pairs could be found mainly from the ORFEUS catalog but also for stations of the other data sets. As displayed in Figure 4.5, the new residuals are on average 0.09 s faster with a standard deviation of 0.77 s for the difference between the residuals. The picking error of the ISC data can then be estimated since the difference in residuals corresponds to the difference in picking errors:

$$d_{ISC} = d + \epsilon_{ISC} \tag{4.1}$$

$$d_{NEW} = d + \epsilon_{NEW} \tag{4.2}$$

$$\Rightarrow d_{ISC} - d_{NEW} = \epsilon_{ISC} - \epsilon_{NEW} \tag{4.3}$$

where d_{ISC} and d_{NEW} denote the individual, observed travel time residuals, d is the residual without picking errors, ϵ_{ISC} and ϵ_{NEW} are the picking errors.
With a picking error on the order of 0.15 s for the newly picked data, a standard deviation of 0.77 s for the distribution of $\epsilon_{ISC} - \epsilon_{NEW}$ measurements and under the assumption of Gaussian error propagation, the picking errors in the ISC data amount to approximately 0.75 s. This value is larger but still in agreement with the estimate of Gudmundsson *et al.* (1990) of $\sigma \approx 0.5$ s for random errors in teleseismic ISC residuals.

4.5 Conclusions

The main objective of this study was to obtain accurate arrival time picks for stations within Europe which did not report to the ISC and to fill data gaps in regions with few stations. Waveforms were provided by spatially dense temporary arrays and data centers in Europe. These waveforms were then preprocessed and absolute arrival times picked with the automatic phase picker of Sandoval *et al.* (2004a). Analyses of the picks show their high quality and that they contain significant information on the geologic properties of crust and uppermost mantle directly beneath the stations. Therefore, the new data set can be combined with the ISC data for global travel time tomography to obtain a high-resolution velocity model of crust and mantle beneath Europe in particular beneath the dense station arrays. Furthermore, this study shows where lack of data/stations is still greatest (e.g. parts of East Europe or Scandinavia). For future research, valuable information could also be gained from near real time picking of the VEBSN (Virtual European Broadband Seismograph Network) as gradually more waveforms become available in near real time for stations which do not report to

the ISC so far.
The new arrival time picks will be made available via anonymous ftp on ftp://terra.geo.uu.nl/-people/amaru/.

4.6 Acknowledgments

We thank ORFEUS, GEOFON, Stephen Arrowsmith, the CALIXTO, EIFEL, MIDSEA, SVEKALAPKO and TOR working groups for providing the data used in this study. This project was funded by ISES (Netherlands Research Centre for Integrated Solid Earth Science).

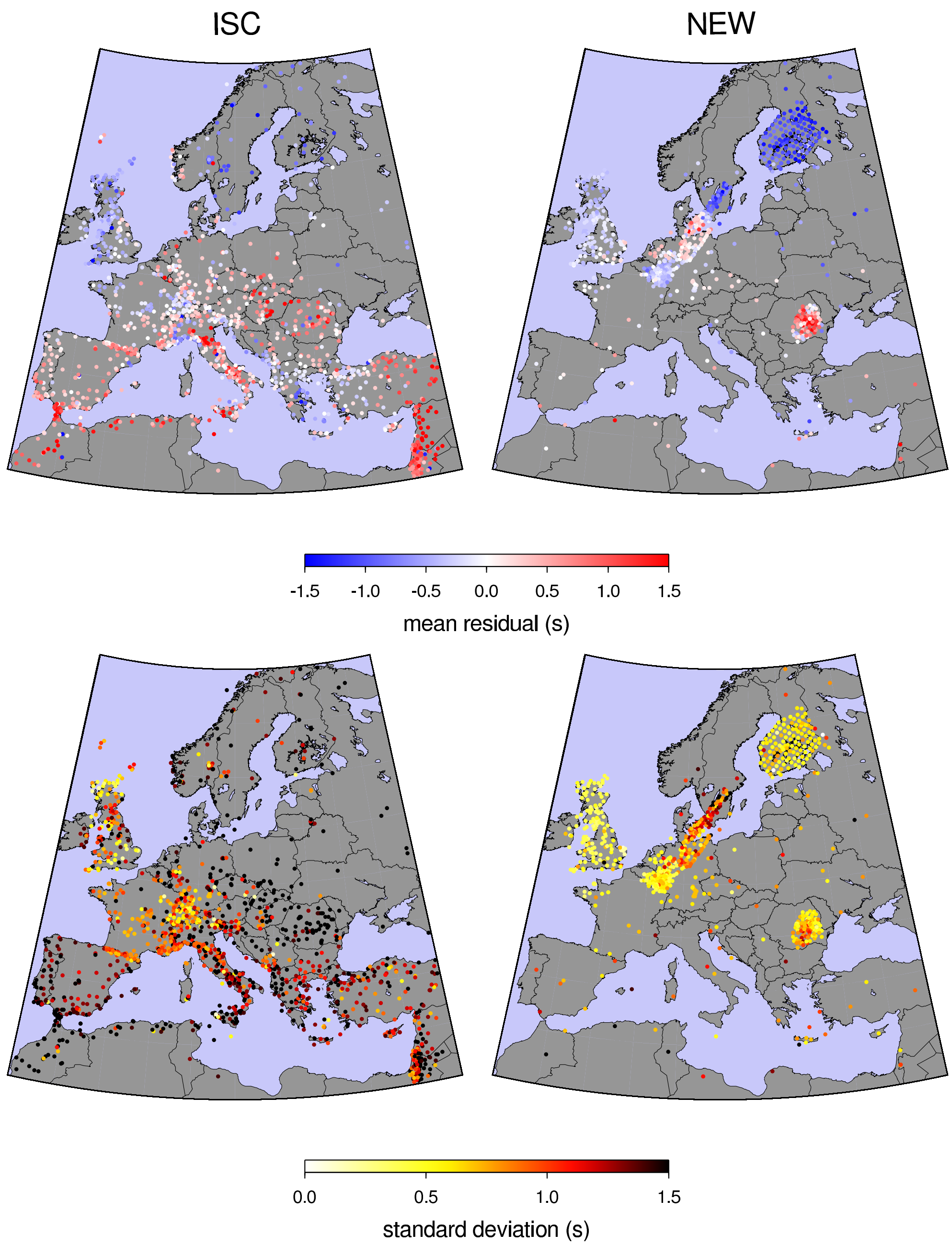

Figure 4.4: Mean teleseismic travel time residuals per station for the EHB data set (upper left) and the new data (upper right) and standard deviation per station for the EHB data set (lower left) and the new data (lower right). All residuals are computed with respect to the reference model ak135.

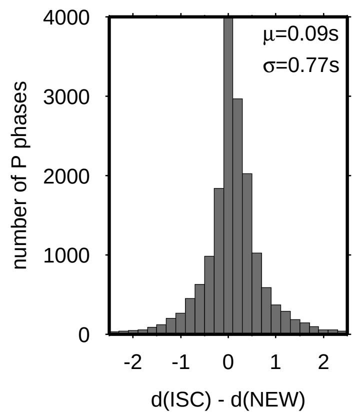

Figure 4.5: Histogram of the differences between residuals of the EHB data set and the newly picked arrival time data set for matching station-event-phase combinations.