Detection of Systematic Errors in Travel-Time Data Using a Minimum 1D Model: Application to Costa Rica Seismic Tomography

by V. Maurer, E. Kissling, S. Husen, and R. Quintero

Abstract   Many seismological studies depend on the accuracy of timing of seismological data. In seismic tomography, travel-time residuals defined as differences between the observed and calculated arrival times of seismic phases are minimized to constrain 3D velocity structure. Inconsistencies and large errors in data sets that result from incorrect station coordinates, errors in the timing acquisition system, errors in the merging procedure, inconsistency in the picking and phase misidentification can also generate travel-time residuals, and because of their systematic nature, these errors cannot be treated as random noise even by exploiting a large number of travel times. While the inverse problem is perfectly set up to deal with random noise, systematic errors lead to significant artifacts in the solution, but may not be detected by a posterior error assessment. For this reason, detecting and removing systematic travel-time errors from data sets before inversion is crucial for seismic tomography studies.

We present a methodology based on the use of a minimum 1D model to detect and remove systematic errors in travel-time data by detailed analysis of station delays and observation residuals and apply it to a local earthquake data set from Costa Rica. The determination of the exact nature of detected inconsistencies needs further investigations in each individual case. If the cause of detected systematic errors cannot be determined beyond any doubt and the afflicted data may not be corrected, they must be deleted from the data set. To assess the extent of influence of systematic errors on hypocenter locations and their uncertainties, we present two examples showing the effects of station mislocation.

Introduction

In most seismological studies, the accuracy of arrival time measurements is crucial to obtain reliable and accurate results. Modern seismic studies frequently benefit from dense station coverage, with higher sampling rates, broader frequency band recorded, and also a higher precision in time and position thanks to continuous GPS records. Unfortunately, the highest possible accuracy is not always reached in practice; data sets still suffer from various errors that can be classified into two distinct and specific groups:

• Random errors,
• Systematic errors.

Large random errors reduce the overall quality of a data set, but their effects can be minimized through the use of a large number of observations and the application of a solution method that correctly addresses Gaussian distributions. Systematic errors, however, which also reduce the quality of a data set, cannot be addressed properly with Gaussian solution methods, because the averaging does not correct for any intrinsic bias, which becomes part of the solution, producing significant artifacts (see, e.g., Kissling, 1988; Röhm *et al.*,

2000). Furthermore, systematic errors are difficult to detect. In studies of routine earthquake location or local earthquake tomography, travel-time residuals are expected to be representative of the difference between the real structure of the studied area and the initial velocity model used in the inverse problem. Inversion algorithms are designed to find systematic interdependent travel-time residuals and attribute them to specific 3D model structure. Systematic errors, however, yield travel-time residuals with similar characteristics as real 3D geologic structure; they result in artifacts in the data that may not be distinguished from real 3D structure. Thus, they lead to serious misinterpretations of seismic tomographic results (Kissling, 1988). Consequently, we must check for systematic errors in the data before the inversion rather than hoping to detect artifacts in the solution.

There are many reasons for systematic errors, including:

1. Incorrect station coordinates,
2. Errors in the acquisition time at a station,
3. Inconsistency in the picking due to noise level changes, different filtering procedures, different seismologists,

4. Phase misidentification (*Pg*, *PmP*, *Pn*),
5. Errors in the merging of data from different networks (e.g., two different events occurring almost at the same time could be misinterpreted as a single event).

Systematic errors in seismic station parameters have been recognized since the earliest routine seismic data collection. In particular, the International Seismological Center (ISC) bulletin exhibits such problems (Grand, 1990; Röhm *et al.*, 1999, 2000). Furthermore, the importance of using accurate and consistent absolute arrival times for local earthquake tomography has been studied in detail recently, concluding that for high resolution tomography and for consistent and accurate hypocenter determination such bulletin data must be repicked (e.g., Diehl *et al.*, 2009). Bulletin data suffer from the limitations that waveforms are not always available and, hence, a method to detect systematic errors in bulletin data is needed that does not require waveforms.

In this article, we propose to use the so-called minimum 1D model (Kissling *et al.*, 1995) to detect and remove systematic errors from local earthquake data. We will see that the minimum 1D model allows us to employ the interdependency of the unknown parameters to detect systematic errors. Indeed, most artifacts may not be introduced into the 1D-velocity model and end up in specific parts of the solution that can be easily checked, such as hypocenter locations or station delays. To illustrate the method, we use examples from a tomography experiment carried out in Costa Rica from 2005 to 2006. In order to avoid problems arising from

picking inconsistencies, arrival times of local earthquakes were determined using an automated quality-weighted picking procedure (Aldersons, 2005; Diehl *et al.*, 2009). We will document the procedure to detect problems at individual stations as well as at sites where two stations from different networks were located nearby.

Local Earthquake Data Set for Costa Rica

The Costa Rica Subduction Modeling (CORISUBMOD) project was set up to better understand subduction-zone related processes and their variations beneath Costa Rica, by improving and extending existing 3D velocity models through local earthquake and teleseismic tomography. In order to reach these goals, 15 broadband sensors (CORISUBMOD network) were installed in Costa Rica for a two-year experiment between 2005 and 2006. Data were continuously recorded during that period. The CORISUBMOD network complemented the two permanent countrywide networks (OVSICORI-UNA and RSN, see Table 1) that have been operated since 1970s to record local and regional seismicity. Seismic stations from other temporary networks and from networks in the neighboring countries (see Table 1) operating in our study area during that time period were also included in our project to build the largest possible data set. Hence, our data set in Costa Rica includes stations from four temporary networks, two countrywide permanent networks, and two permanent networks in neighboring countries, combining a

Table 1

Label, Reference, and Contact for the Networks

Network Label	Institution and Contact
CORISUBMOD	Costa Rica Subduction Modeling (CORISUBMOD) a temporary seismic network jointly deployed by the Observatorio Vulcanológico y Sismológico de Costa Rica (OVSIRORI-UNA), by the Red Sismológica Nacional (RSN) and by the ETH Zurich, Switzerland in Costa Rica from 2005 to 2007. ETHZ, Institute of Geophysics, Sonneggstr. 5. CH-8093 Zurich, Switzerland. Contact: husen@sed.ethz.ch
Kiel	Surface and borehole seismic stations administrated by the University of Kiel in central and northern Costa Rica from 2005 to 2006. CAU Kiel—IFG—Geophysik, Otto-Hahn-Platz 1, 24118 Kiel, Germany. Contact: wrabbel@geophysik.unikiel.de
OVSICORI-UNA	Observatorio Vulcanológico y Sismológico de Costa Rica, administrated by the Universidad Nacional de Costa Rica (OVSICORI-UNA). Apartado Postal: 2346-3000 Heredia, Costa Rica. Contact: rquinter@una.ac.cr or http://www.ovsicori.una.ac.cr
OSIVAM	Observatorio Sismolórico y Vulcanolórico Arenal-Miravalles (OSIVAM), administrated by the Instituto Costarricense de Electricidad (ICE), Apdo. Postal 10032-1000 San José, Costa Rica. Contact: WTaylor@ice.go.cr or http://www.arenal.net/volcano-eruption/arenal-miravalles-seismological.htm
PANAMA	Instituto de Geociencias, Red Sismológica National, Campus Central-COLINA, Universidad de Panamá Ciudad de Panamá, administrated by Universidad de Panamá. Contact: ecamacho@cableonda.net or http://www.igc.up.ac.pa/
Pocosol	Pocosol Seismic Network, an NSF-funded array of three stations deployed by Rutgers University in northern Costa Rica from 2005 to 2006. Earth and Planetary Sciences, Rutgers University, Wright-Rieman Labs, 610 Taylor Road, Piscataway, New Jersey 08854. Contact: vlevin@rci.rutgers.edu or linkimer@email.arizona.edu
RSN	Red Sismológica Nacional (RSN) administrated jointly by the Instituto Costarricense de Electricidad (ICE) and the Universidad de Costa Rica (UCR), Escuela Centroamericana de Geología, Apdo. Postal 214-2060, San José, Costa Rica. Contact: credondo@geologia.ucr.ac.cr or http://www.rsn.geologia.ucr.ac.cr/
TUCAN	Tomography Under Costa Rica and Nicaragua (TUCAN) is a temporary network deployed by Boston University in Nicaragua and northern Costa Rica from 2004 to 2006. Boston University, 675 Commonwealth Ave., Boston, Massachusetts 02215. Contact: abers@bu.eduAbers or http://people.bu.edu/abers/TUCAN.html

set of 135 seismic stations (see Figure 1 for locations and Table 1 for references). The stations were equipped with one-component short-period sensors (51), three-component short-period sensors (30), and broadband sensors (54). To avoid inconsistencies in the picking procedure and reduce random errors, we used an automatic picker called Manneken-PIX (MPX) (Aldersons, 2004), which, among the arrival time pick, also provides a corresponding quality estimate, which is related to the timing error of the arrival time. This information allows us to assess the overall quality (timing error) of our data set, a key information and prerequisite for seismic tomography (see, e.g., Diehl *et al.*, 2009). The MPX system combines a traditional picking algorithm based on Baer and Kradolfer (1987) with an automatic quality assessment, which is able to imitate the picking behavior performed by experienced seismologists. For details about the working flow of MPX see Diehl *et al.* (2009). The picking engine of MPX needs to be guided by *a priori* information of the appropriate arrival time for each seismogram. It can either be a routine pick or a predicted time computed from a reasonable velocity model. For our study, we used only predicted picks, as routine picks were not available for temporary networks. A previously calculated so-called minimum 1D model (Quintero and Kissling, 2001) was used to compute predicted arrival times to guide MPX.

To assess the quality (timing error) of each observation, MPX uses a pattern recognition scheme that needs to be tuned to a handpicked reference data set (see Di Stefano *et al.*, 2006; Diehl *et al.*, 2009). A set of 1158 reference seismograms from a representative selection of 27 reference events has been carefully handpicked following the consistent picking rules given by Diehl *et al.* (2009). The reference hand-

picked observations were set up as basis to calibrate MPX in order to automatically repick 1444 events (31,863 waveforms) reported in 2005 and 2006. A final data set of 764 well-locatable events (11,081 observations) was used to compute an updated minimum 1D model (later referred to as CR_min1D2007) for the CORISUBMOD data set. The updated minimum 1D model did not show any significant differences to the minimum 1D model of Quintero and Kissling (2001), indicating that a minimum in the 1D-velocity model and in the station delays had already been reached. All events were relocated with VELEST (Kissling *et al.*, 1995) in single-event mode, using the CR_min1D2007 model. Only events having a maximum azimuthal gap of less than 175° degrees, and having six or more observations, were used in our study.

Detecting Systematic Errors: Approach

The travel time of a seismic wave is a nonlinear function of the hypocentral parameters and seismic velocities sampled along the ray path between source and station. This dependency is called the coupled hypocenter-velocity problem that can be iteratively solved by standard local earthquake tomography procedures for a large number of hypocenters (e.g., SIMULPS; Thurber, 1992). In standard earthquake location procedure, the velocity parameters are kept to *a priori* values and the observed travel times are interpreted by perturbations of the hypocentral parameters only. Neglecting the coupling between hypocentral and velocity parameters during the location process can introduce systematic errors in hypocenter locations (Thurber, 1992). Furthermore, derived location uncertainties strongly depend on the assumed *a priori* velocity structure and normally largely underestimate the true location errors (Kradolfer, 1989; Husen *et al.*, 1999). Precise hypocenter location and, in particular, their uncertainty estimates, therefore, demand the simultaneous inversion of velocity and hypocenter parameters. The Shareware program VELEST (Kissling, 1998) iteratively solves the coupled hypocenter-velocity problem for a large number of events and a 1D velocity model with station delays. Although this program has originally been developed to derive a well-suited initial reference velocity model for 3D local earthquake tomography (Kissling *et al.*, 1994), it may also be applied to the joint hypocenter determination problem to allow quantitative uncertainty estimates. The so-called minimum 1D velocity model (for a detailed definition of the term and for a calculation guide of a minimum 1D velocity model, see Kissling *et al.*, 1994, 1995, respectively) is obtained by a trial-and-error process using VELEST with various initial velocity models and hypocentral parameters and for different damping values. It represents a model that leads to a minimum in average values for all earthquakes and closely reflects the *a priori* structural information obtained from surface geology and/or by controlled source studies (for Costa Rica, see Stavenhagen *et al.*, 1997 and Sallarès *et al.*, 2001). To account for lateral variations in the shallow
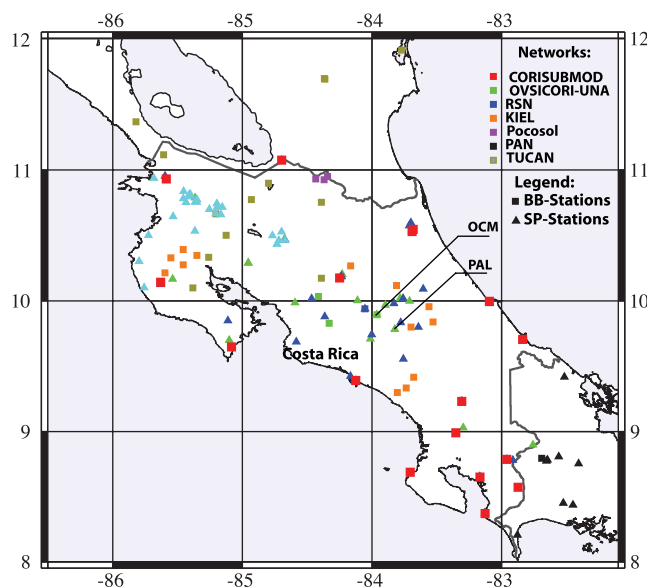

Figure 1. Locations of seismic stations used in the CORISUB-MOD project, sorted by networks and type of sensor, as indicated on the inset. Stations PAL and OCM are also identified.

subsurface, station delays are incorporated in the inversion process. Because results of this inverse problem are ambiguous, different velocity models with similar residual variance can be obtained. The one that most closely reflects the *a priori* known information about the near-surface structure and that leads to a minimum average RMS value for all earthquake locations is selected as the reference model.

The minimum 1D model solution to the coupled hypocenter-velocity problem is basically composed of a 1D velocity model, a set of hypocenter locations, and a station delay attributed to each station. All these parameters are closely linked. The minimum 1D model solution reflects their interdependencies and allows us to check each parameter individually, because unlike a 3D model, a 1D model may not absorb most systematic errors with false 3D structures (artifacts) because the velocity part is highly overdetermined. As each of the unknown parameters are close to a minimum with regard to the specific coupled inverse problem, the minimum 1D model solution allows us to detect systematic errors in a data set by focusing on one parameter at a time. It should be noted here that the main use of the minimum 1D model in this article is not to obtain the most accurate hypocenter locations, but to derive a set of hypocenter locations, station delays, and velocity parameters that allow the method to detect systematic data errors.

Minimum 1D Model for Costa Rica

We propose to use the minimum 1D model solution to detect systematic errors at stations. For obvious reasons, we assume that all observations from stations with *a priori* known and unrecoverable problems are previously removed from the data set. We further assume that not all stations will suffer from the same systematic errors. Hence, a significant data set from a large number of stations will provide reliable background information which will constitute the basis of our approach.

The first minimum 1D model for Costa Rica was calculated by Quintero and Kissling (2001). In their study, they explained the procedure to merge events recorded by two permanent networks in Costa Rica (RSN and OVSICORI-UNA; see Table 1). Because many of the stations used in the CORISUBMOD experiment were not operational at the time of the Quintero and Kissling (2001) study, we had to update their minimum 1D model for our data set. For these calculations, we considered only events with an azimuthal gap smaller than 180° and with a number of observations greater than six. This assures us that the hypocentral parameters can be well determined. The updated minimum 1D velocity model CR_min1D2007 is virtually identical with the one calculated by Quintero and Kissling (2001). We obtained station delays for the new stations; a comparison showed us that station delays for the older sites obtained for our new minimum-1D model were virtually the same as those obtained in the original study (Quintero and Kissling, 2001).
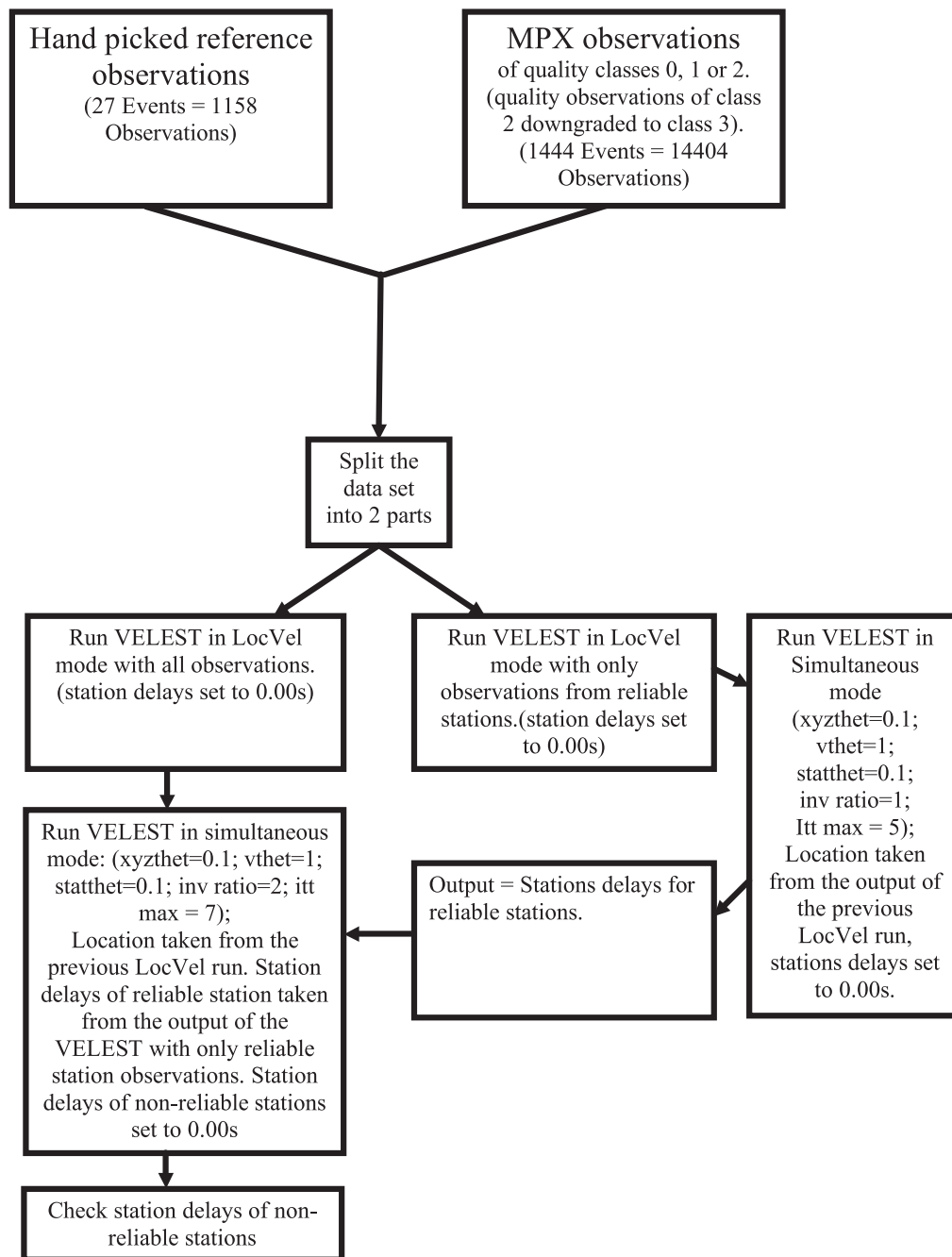
A list of potentially problematic stations was compiled during the calculation of CR_min1D2007 based on the analysis of the VELEST output. The list also contained stations suspected as potentially problematic based on *a priori* knowledge acquired when compiling and handpicking the station list. Stations may be flagged as potentially unreliable for a variety of reasons, including (but not limited to) *a priori* knowledge such as coordinate issues, known timing problems, observations during phase picking; or they may be flagged due to VELEST results such as the regular appearance of a station in earthquakes with a large final RMS or suspicious azimuthally dependent or unrealistic station delays compared to near-surface geologic features.

Analysis of Results Obtained by Data from Potentially Problematic Stations

Once the list of stations that needed to be checked was constituted, the original data set was temporarily reduced by extracting and deleting all observations from questionable stations. With the remaining data set containing only observations from likely reliable stations and using the previously calculated minimum 1D model CR_min1D2007 and hypocenter locations as initial input, a set of station delays for likely reliable stations was calculated. These station delays were then input in a second run of VELEST, using at this time the original data set with all observations. The initial station delays of suspected unreliable stations were set to zero in the second run. The purpose of this procedure was to enhance inconsistent station delays or travel-time residuals at unreliable stations. These parameters depended strongly on the data from potentially problematic stations. Indeed, many systematic errors in travel-time data thus appeared once they were compared in the minimum 1D model solution with the reliable data. Figure 2 illustrates with more details the methodology designed for this study.

Detecting Systematic Errors: Application to Costa Rica

When a systematic error such as wrong station coordinates or a defective timing acquisition system of a station is suspected, the usual procedure is to send someone to check it in the field. However, this check is sometimes not possible, because of logistics or if the temporary station has already been dismantled. Data from any seismic station is precious; hence, another way to confirm its reliability must be devised. Here we apply our approach to local earthquake data from Costa Rica. Because the CORISUBMOD data set consists of a mixture of temporary and permanent stations operated by different agencies, inconsistencies in error assessment and even errors in station parameters are likely expected. Some of these problems were already detected during the handpicking of the reference set for subsequent automated picking. As we aim for a consistent phase picking for all waveforms, we

```
┌─────────────────────────┐        ┌─────────────────────────┐
│  Hand picked reference  │        │    MPX observations     │
│      observations       │        │ of quality classes 0, 1 or 2. │
│   (27 Events = 1158     │        │ (quality observations of class │
│      Observations)      │        │  2 downgraded to class 3). │
│                         │        │   (1444 Events = 14404  │
│                         │        │      Observations)      │
└─────────────────────────┘        └─────────────────────────┘
```

```
                    ┌──────────────┐
                    │  Split the   │
                    │  data set    │
                    │ into 2 parts │
                    └──────────────┘
```

```
┌─────────────────────┐  ┌─────────────────────┐  ┌─────────────────┐
│ Run VELEST in LocVel│  │ Run VELEST in LocVel│  │  Run VELEST in  │
│ mode with all obser-│  │ mode with only      │  │   Simultaneous  │
│ vations.            │  │ observations from   │  │      mode       │
│ (station delays set │  │ reliable            │  │  (xyzthet=0.1;  │
│ to 0.00s)           │  │ stations.(station   │  │     vthet=1;    │
│                     │  │ delays set          │  │   statthet=0.1; │
│                     │  │ to 0.00s)           │  │   inv ratio=1;  │
│                     │  │                     │  │   Itt max = 5); │
└─────────────────────┘  └─────────────────────┘  │ Location taken  │
                                                   │ from the output of │
┌─────────────────────┐                            │ the previous    │
│ Run VELEST in simul-│  ┌─────────────────────┐   │    LocVel run,  │
│ taneous mode: (xyz- │  │ Output = Stations   │   │ stations delays set │
│ thet=0.1; vthet=1;  │  │ delays for reliable │   │    to 0.00s.    │
│ statthet=0.1; inv   │  │ stations.           │   └─────────────────┘
│ ratio=2; itt        │  │                     │
│ max = 7);           │  │                     │
│ Location taken from │  └─────────────────────┘
│ the previous LocVel │
│ run. Station delays │
│ of reliable station │
│ taken from the      │
│ output of the       │
│ VELEST with only    │
│ reliable station    │
│ observations. Station│
│ delays of non-reliable│
│ stations set to 0.00s│
└─────────────────────┘

┌─────────────────────┐
│ Check station delays│
│ of non-reliable     │
│ stations            │
└─────────────────────┘
```

*All VELEST runs were computed with CRmin1D2007 velocity model.

Figure 2. Procedure to detect stations with systematic errors in travel-time data with VELEST.

strive to assess the quality of the station parameters systematically for all stations.

Detecting Wrong Station Coordinates Using Travel-Time Residuals: Example of Station PAL

The problem of wrong station coordinates is well-known and the analysis of travel-time residuals as a function of azimuth is often used to detect wrong stations coordinates.

Consequently, we first demonstrate the performance of our approach by studying this well-known problem. After relocation of all hypocenters and calculation of station delays with the minimum 1D model CR_min1D2007, observations from station PAL (equipped with a 1-component, short-period sensor) showed large residuals. For this reason, this station was placed in the list of potentially unreliable stations and the previously described procedure was applied. Figure 3 is a polar diagram showing the travel-time residuals of PAL
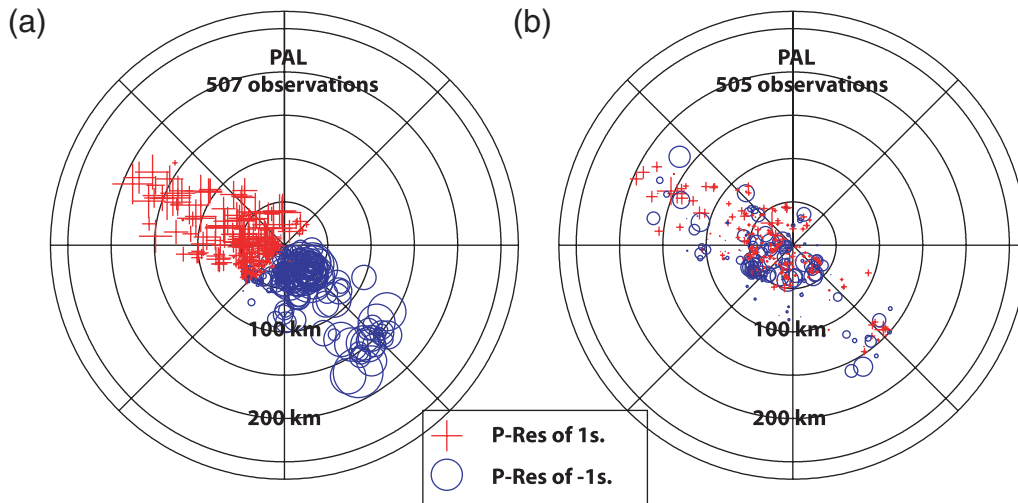
Figure 3. Polar diagram distance/azimuth showing the travel-time residuals in seconds at station PAL: (a) using wrong station coordinates, and (b) using correct station coordinates.

as a function of distance and azimuth. The residuals shown were calculated using the minimum 1D model and using station delays of reliable stations. Figure 3a shows a bipolar distribution of the travel-time residuals at PAL, with all the positive values situated in the northwest quadrant and all the negative travel-time residuals in the southeast quadrant relative to the station. We further note that the amplitude of the positive and negative residuals are not distance-dependent, with an almost constant value of 1 sec. Two different interpretations are possible. Either PAL is exactly placed at the border of a very large change in the near-surface geological structure, or the station is mislocated. Station PAL is indeed situated near the limits of quaternary volcanic rocks to the north and a small sedimentary basin to the south. But the latter is itself situated on top of a series of Miocene–Pliocene volcanic rocks (Tournon and Alvarado, 1997). The small size of the sedimentary basin, however, could not explain the amplitude of the residuals observed. Moreover, even if the station is placed at the border of a significant structural change, the residuals should show some increase with increasing distance of the sources. Furthermore, it is unlikely that such a geologic structure would result in positive and negative residuals of the same amplitude. For these reasons, the pattern shown by Figure 3 is interpreted as a mistake in the coordinates of the station and we test this hypothesis by estimating the likely correct coordinates. The travel-time residual denotes the difference between the calculated travel time and the observed travel time. With the station delay we observe an average travel-time residual of about 1 sec. Therefore, if we consider an average crustal *P*-wave velocity of 6 km/sec, the station PAL should be located around 6 km to the southeast from its current coordinates. Based on our inquires the coordinates of PAL were checked in the field and, indeed, a mistake was found. The real location of the station is 7.3 km southeast (see Table 2), in good agreement with our estimates. Once relo-

cation is done with the correct coordinates, the bipolar distribution no longer appears (see Fig. 3b) and travel-time residuals are now dependent on epicentral distance.

Detecting Systematic Errors in Travel-Time Data Using the Station Delay Value

The seismic station crt06 is a broadband sensor from a temporary network (see Fig. 4). We were warned by the operators that all stations of this temporary network could suffer from timing errors because of a problem in the quartz crystals of the timing system in the data loggers. For this reason, all stations from this temporary network were added to the list of potentially unreliable stations. Another example was detected during handpicking of arrival times at stations CDM and BUS. Arrival times at station BUS showed large differences when compared with arrival times picked at station CDM, situated only 820 meters away. For this reason BUS and CDM were also added to the list of unreliable stations. Figure 4 shows the station delays at several stations situated near crt06, BUS, and CDM in relation to the station delays obtained for the three of them with the previously described procedure. Stations crt03, crt04, and crt05, together with crt06 form a transect from the coast to the Talamanca cordillera. Their station delays are, respectively,

Table 2
PAL Coordinates

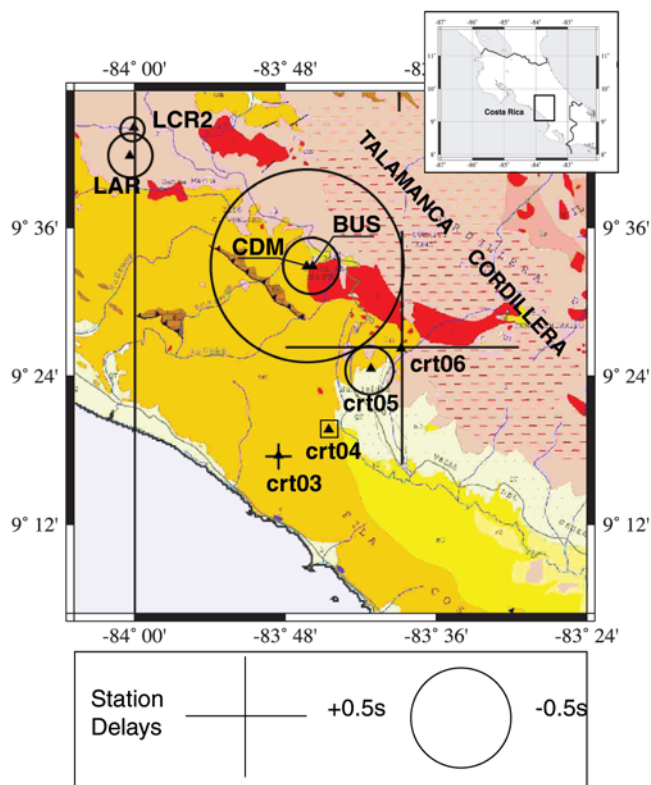| PAL	Wrong Coordinates	Correct Coordinates
Latitude	9.8295	9.7830
Longitude	−83.8720	−83.8240
Altitude	1220	1352
Distance	7.373 km	

Figure 4. Station delays at crt03, crt04, crt05, crt06, BUS, CDM, LAR, LCR2, and geological setting. Yellow colors represent sediment rocks; red colors represent igneous rocks. (Geological map taken from Tournon and Alvarado, 1997.)

0.11 sec, 0.02 sec, and −0.20 sec seconds, and 0.98 sec for crt06 (see Table 3). Except for crt06, these delays are in good agreement with a reduction in the thickness of a low-velocity Oligocen-Miocen sedimentary basin and the limits of the region of high-velocity, volcanic postintrusive rocks of the Talamanca cordillera along this transect (Tournon and Alvarado, 1997). Not only is the sign of the station delay calculated at crt06 in disagreement with the high velocities associated with volcanic rocks underneath the station, but also the large value should correspond to a low-velocity zone such as a deep sedimentary basin. Hence, the station delay allowed us to confirm a problem in the timing acquisition at crt06. Unfortunately, the timing problems potentially varied over the period of operation of the station and no time correction could be defined. Consequently, all observations from this station were deleted from our data set.

The station delay found at BUS from 479 observations is −0.19 sec, while at CDM a station delay of −0.81 sec with 21 observations has been computed (see Table 3). Despite the difference in number of observations, such a big difference in station corrections is suspicious considering that the two stations are situated in the same geological environment (dioritic rock sequences; Tournon and Alvarado, 1997). The velocity of the subsurface rocks nicely corresponds with the delay obtained for BUS, while the value calculated at CDM is unlikely large and showing an unlikely large azimuthal variation. Unfortunately, the timing accuracy at station CDM could not be confirmed in the field, leaving the large station residual unexplained. For these reasons, all observations from CDM were removed from the data set.

Transient Timing System Problems

Station OCM is composed of two sensors: a one-component, short-period sensor (OCM) and a three-component, broadband sensor (OCMB). The most intuitive test to check the timing acquisition at both sensors is to compare the difference between arrival times determined at each sensor for the same events. Figure 5 illustrates the calculated differences between OCMB and OCM in our study period (2005 to 2006). During most of the observation period (period 1a from 1 January 2005 to 29 July 2005 and period 1b from 17 January 2006 to 31 December 2006) the differences are nearly zero, while from 29 July 2005 to 17 January 2006 (period 2) the calculation shows a difference in arrival times close to 1 sec. Individual differences greater than ±0.2 sec in periods 1a and 1b were all visually controlled; it was found that they all relate to low quality emergent onsets with rather large uncertainties. A difference of 1 sec in period 2, however, obviously indicates a problem in the time acquisition system by at least one of the sensors. But by calculating the difference and inspecting the delays, we are not able to identify which sensor is malfunctioning. The situation could be even worse: Another possibility is that both sensors were incorrect during periods 1a and 1b and one is correct only during period 2. Without additional information and in order to be consistent and to keep a clean data set, we would have to remove all observations from both sensors. In this section we will see that by using VELEST with a minimum 1D model, it is possible to identify which sensor was incorrect and during which time period.

To solve the problem at OCM, we calculated station delays and residuals for the broadband and the short-period sensors for the time periods 1a, 1b, and 2 separately (see

Table 3

Station Delays at crt03, crt04 crt05 crt06, BUS, CDM, LAR, and LCR2

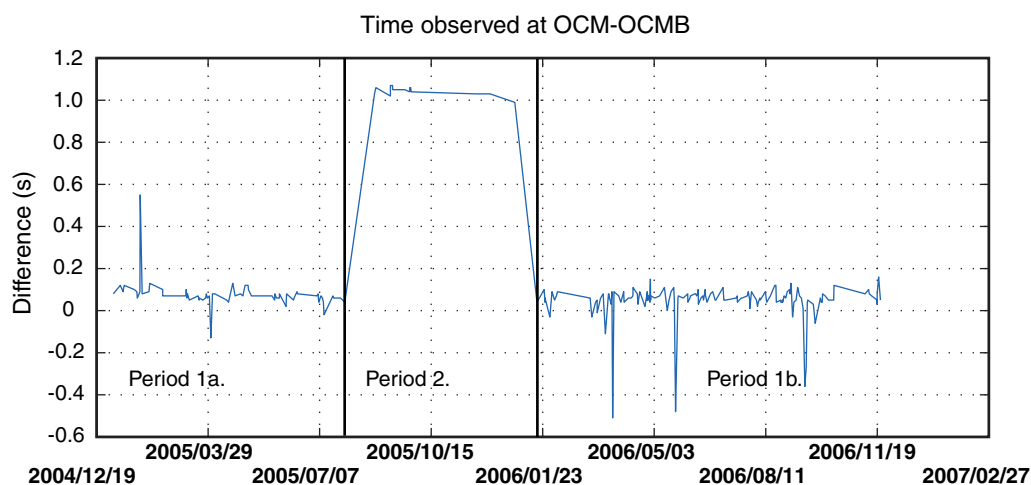| Stations	crt03	crt04	crt05	crt06	CDM	BUS	LAR	LCR2
Observations	24	72	141	77	21	479	526	506
Delays	0.11	0.02	−0.20	0.98	−0.98	−0.19	−0.18	−0.13

Time observed at OCM-OCMB

Figure 5. Difference between arrival time observations recorded at the two sensors of station OCM.

Fig. 5 for definition of time periods 1a, 1b, and 2). We used VELEST in simultaneous mode with all travel-time data, and as input, the minimum 1D velocity model *CR_min2007* with station delays except for stations OCMB and OCM, which were set to zero. Table 4 displays the results.

For time periods 1a and 1b combined, the station delays were -0.18 sec at OCMB and -0.14 sec at OCM, whereas for time period 2 the station delays were -1.16 sec at OCMB and -0.16 sec at OCM. Hence, we note that the station delay found at OCMB for period 2 is about 1 sec lower than the delay found for periods 1a and 1b; this corresponds well with the difference between the travel-time arrival at OCM and OCMB during period 2. This clearly demonstrates that the broadband sensor OCMB suffered from a timing problem from 29 July 2005 to 17 January 2006. For this reason, all observations from OCMB for period 2 were removed from our data set.

Discussion

As illustrated by the examples presented here, using a minimum 1D model solution allows us to detect systematic errors as small as 1 sec. In subduction zones, such as Costa Rica, a delay of 1 sec is not unrealistically large and, hence, may not be easily detected in the data. A systematic error of 1 sec at one station will certainly introduce a large local velocity variation, but this anomaly will not appear as an obvious artifact in comparison with the velocity variations normally expected in complex regions such as subduction

zones, which can be easily in the order of 10% (see, e.g., Husen *et al.*, 2003). In summary, systematic errors such as the ones documented in this article will likely go undetected in the data and the results. The minimum 1D model solution, however, allows us to detect and remove such systematic errors in travel-time data, thus avoiding artifacts and possible misinterpretations in seismic tomography results. The advantage of the minimum 1D model solution lies in the simultaneous inversion of seismic velocities, hypocenter locations, and station delays. The use of a 1D velocity model ensures that the effect of systematic data errors will most likely show up in the hypocenter parameters and in the station delays, because the velocity part is highly overdetermined. This allows a more reliable detection of systematic data errors as demonstrated in this article. The proposed procedure is even more straightforward in cases where the computation of a minimum 1D model is planned as part of a local earthquake tomography study.

The effect of a station with wrong coordinates on earthquake location is usually estimated to be rather small because there are many observations and only four unknowns in this inverse problem. To have an idea of the effect of station mislocations on earthquake hypocenter locations, we used two events recorded by station PAL as an example. We computed synthetic arrival times for all stations that recorded those two events using the 3D velocity model calculated by Husen *et al.* (2003). To be more realistic, we introduced normal distributed Gaussian noise with a variance equal to the quality class

Table 4
Stations Delays Obtained at OCM for Periods 2 and $1a + 1b$

	Period 2		Period $1a + 1b$	
	OCM	OCMB	OCM	OCMB
Number of Observations	45	24	329	253
Station Delays	-0.18 sec	-1.16 sec	-0.18 sec	-0.14 sec

(uncertainty) of the real observations. The two events were then iteratively relocated using synthetic arrival times from stations that originally recorded the event and for different locations of PAL. To estimate the precision of the relocated hypocenter location, we calculated the hypocenter error ellipsoid for every kilometer of station mislocation. The error ellipsoids are computed from the nonlinear probabilistic solution of the earthquake location problem by the shareware NonLinLoc (Lomax *et al.*, 2000; Husen *et al.*, 2003). This is an important point since NonLinLoc computes the full set of uncertainties, which are more reliable than uncertainties computed by traditional, linearized methods. As in the pre-

viously described example with real data, PAL coordinates were shifted to the northwest.

Figure 6 illustrates the effect introduced by wrong station coordinates on the hypocentral parameters. Obviously, the geometry of the observing stations plays an important role in the location. In Figure 6a, PAL is just one of several stations covering a small azimuthal sector and distance range. The same calculations were repeated (see Fig. 6b), but removing the two more distant stations in the azimuthal sector of PAL thus increased the importance of the observation at PAL. In the second case, a mislocation of 7 km in the station coordinates, such as it was found for PAL, yields a
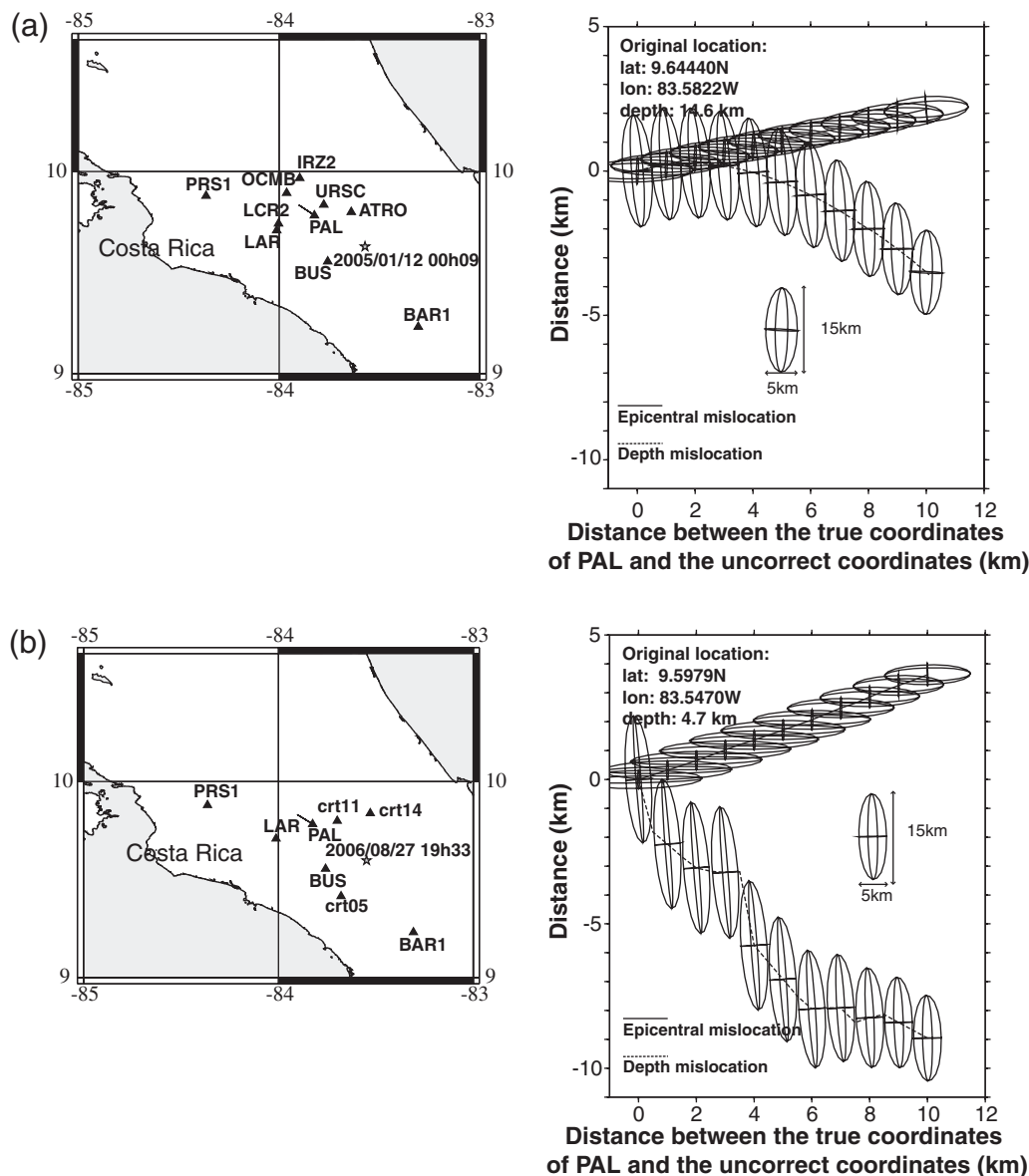

Figure 6. Effects of wrong station coordinates on earthquake location. Maps on the left side show the earthquakes epicenter locations. The solid lines on the maps represent the successive uncorrect coordinates of PAL. On the right side, the solid lines represent the horizontal distance between the real earthquake location and the earthquake location calculated with wrong coordinates for station PAL. The dashed lines represent the vertical difference. The error ellipsoids are reduced by a factor 5.

mislocation of 2.5 km in the epicenter location and 8 km in focal depth (Fig. 6b). While the horizontal error increases linearly with increasing station mislocation, the effect on focal depth is much more severe and erratic in function of the geometry of the observing stations.

More important than the increase in mislocation, however, is that the size of the half axes of the error ellipsoids is not varying at all in function of the coordinate shift. This means that hypocenter uncertainty estimates completely neglect this systematic error. The error in station coordinates is absorbed in the earthquake location; there would be absolutely no way to detect the wrong station coordinates by looking at the estimated error calculated with the hypocenter locations.

Conclusions

This study has taken a step in the direction of detecting systematic errors in local earthquake data. Systematic errors in travel time lead to severe artifacts and potentially misinterpretation of tomographic results. As these errors are systematic, they cannot be compensated by using a large number of observations. Their effects on seismic tomography results have already been shown in previous studies (see, e.g., Kissling, 1998; Röhm et al., 2000). In this article, we focus on detecting such errors by the analysis of large travel-time data sets. To detect and remove systematic errors, we set up a methodology using the program VELEST to calculate a minimum 1D model solution for the study area. We illustrated how, using station delays and/or travel-time residuals, the effect of systematic errors becomes isolated and thus visible. The velocity model plays an important role in the detection of systematic errors, but has to be close to a minimum to distinguish systematic errors from random errors and geologic information.

We clearly demonstrated, with the study on the effect of systematic errors on hypocenter locations, the complexity of the detection of such errors in the solution and in the error estimates. This was illustrated by the invariant size of the half axes of the error ellipsoids. Their effects on hypocenter locations are, however, not negligible and could be comparable to studies on effects of mispicks on earthquake location (Shearer, 1997).

Systematic errors have to be removed from any data set in order to avoid misinterpreting results and to reach the highest possible accuracy. Our procedure allows for important data quality checks and the detection of systematic errors in cases where the computation of a minimum 1D model is planned as part of a 3D inversion.

Data and Resources

Waveforms used in this study were collected from different permanent and temporary networks installed in Costa Rica between 2005 and 2006. Data can be obtained by contacting the person in charge of each individual network,

listed in Table 1. Some plots were made using the Generic Mapping Tools version 4.2.1 (www.soest.hawaii.edu/gmt; Wessel and Smith, 1998).

References

Aldersons, F. (2004). Toward three-dimensional crustal structure of the Dead Sea region from local earthquake tomography, *Ph.D. Thesis*, Tel Aviv University, Israel.

Aldersons, F. (2005). *MannekenPIX 1.7—User Guide*, http://faldersons .net.

Baer, M., and U. Kradolfer (1987). An automatic phase picker for local and teleseismic events, *Bull. Seismol. Soc. Am.* **77,** 1437–1445.

Di Stefano, R., F. Aldersons, E. Kissling, P. Baccheschi, C. Chiarabba, and D. Giardini (2006). Automatic seismic phase picking and consistent observation error assessment: Application to the Italian seismicity, *Geophys. J. Int.* **165,** 121–134.

Diehl, T., E. Kissling, S. Husen, and F. Alderson (2009). Consistent phase picking for regional tomography models: Application to the greater alpine region, *Geophys. J. Int.* **176,** 542–554.

Grand, S. (1990). A possible station bias in travel time measurements reported to ISC, *Geophys. Res. Lett.* **17,** no. 1, 17–20.

Husen, S., E. Kissling, E. Flueh, and G. Asch (1999). Accurate hypocenter determination in the seismogenic zone of the subducting Nazca plate in northern Chile using a combined on-/offshore network, *Geophys. J. Int.* **138,** 687–701.

Husen, S., R. Quintero, E. Kissling, and B. Hacker (2003). Subduction-zone structure and magmatic processes beneath Costa Rica constrained by local earthquake tomography and petrological modelling, *Geophys. J. Int.* **155,** 11–32.

Kissling, E. (1988). Geotomography with local earthquake data, *Rev. Geophys.* **26,** 659–698.

Kissling, E., W. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer (1994). Initial reference models in local earthquake tomography, *J. Geophys. Res.* **99,** 19,635–19,646.

Kissling, E., U. Kradolfer, and H. Maurer (1995). VELEST user's guide-short introduction, *Tech. rep.*, Institute of Geophysics, ETH Zurich, kiss@tomo.ig.erdw.ethz.ch.

Kradolfer, U. (1989). Seismiche Tomographie in der Schweiz mittels lokaler Erdbeben, *Ph.D. Thesis*, Institute of Geophysics, ETH Zurich.

Lomax, A., J. Virieux, V. P. , and C. Berge (2000). Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations, in *Advances in Seismic Event Location*, Kluwer, Amsterdam, pp. 101–134.

Quintero, R., and E. Kissling (2001). An improved *P*-wave velocity reference model for Costa Rica, *Geofisica Internacional* **40,** no. 1, 3–19.

Röhm, A., H. Bijwaard, W. Spakman, and J. Trampert (2000). Effects of arrival time errors on traveltime tomography, *Geophys. J. Int.* **142,** 270–276.

Röhm, A., J. Trampert, H. Paulssen, and R. Snieder (1999). Bias in reported seismic arrival times deduced from the ISC bulletin, *Geophys. J. Int.* **137,** 163–174.

Sallarès, V., J. Dañobeitia, and E. Flueh (2001). Lithospheric structure of the Costa Rican Isthmus: Effect of the subduction zone magmatism on an oceanic plateau, *J. Geophys. Res.* **106,** 621–643.

Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence, *J. Geophys. Res.* **102,** no. B4, 8269–8283.

Stavenhagen, A., E. Flueh, C. Ranero, K. McIntosh, T. Shipley, G. Leandro, A. Schulze, and J. Dañobeita (1997). Seismic wide-angle investigations in Costa Rica a crustal velocity model from the Pacific to the Caribbean coast, *Zbl. Geol. Palont. H.* **3–6,** 393–408.

Thurber, C. (1992). Hypocenter-velocity structure coupling in local earthquake tomography, *Phys. Earth Planet. Inter.* **75,** 55–62.

Tournon, J., and G. E. Alvarado (1997). *Mapa geolgico de Costa Rica: Folleto explicativo*, 79 pp., C. R. Tecnolgica de, Costa Rica (in Spanish).

Wessel, P., and W. Smith (1998). New, improved version of the Generic Mapping Tools released, *EOS Trans. AGU* **79,** 579.

Institute of Geophysics
ETH Zurich, Sonneggstr. 5
CH-8093 Zurich, Switzerland
maurer@tomo.ig.erdw.ethz.ch
(V.M., E.K., S.H.)

Observatorio Vulcanologico y Sismologico de Costa Rica
2346-3000 Heredia
Costa Rica
(R.Q.)