Title
Accuracy of modem global earthquake catalogs

Journal
Physics of the Earth and Planetary Interiors, 135(2-3)

ISSN
0031-9201

Author
Kagan, Yan Y

Publication Date
2003-02-01

DOI
10.1016/S0031-9201(02)00214-5

Peer reviewed

PEPI manuscript KC002-04, PEPI 4179; Journal # 5357. Printed: August 20, 2007

ACCURACY OF MODERN GLOBAL EARTHQUAKE CATALOGS

Yan Y. Kagan

Department of Earth and Space Sciences, University of California,
Los Angeles, California 90095-1567, USA:
Email – ykagan@ucla.edu, Tel. 310-206-5611; FAX 310-825-2779

Abstract. We compare several modern (1977-present) worldwide earthquake catalogs to infer their completeness, earthquake origin time and hypocenter location accuracy, magnitude/scalar seismic moment errors, and difference between individual focal mechanism/moment tensor solutions. The Harvard centroid moment tensor (CMT), U.S. Geological Survey (USGS) MT, USGS first-motion (FM) focal mechanism, PDE and ISC catalogs have been analyzed and compared. The catalogs' completeness and accuracy vary in time and depend on earthquake depth and tectonic environment. We propose a new statistical method for evaluating catalog completeness and show the results for the CMT dataset. A difference in frequency range of seismic waves used in earthquake processing leads to varying degrees of catalog completeness for foreshocks and aftershocks close in time. Earthquake origin time *versus* centroid time as well as hypocenter location *versus* centroid location can be explained well by earthquake scaling relations. Comparing moment magnitudes and regular earthquake magnitudes yields estimated magnitude uncertainties and shows that latter magnitudes poorly estimate earthquake size for large events. Moment errors reported in the CMT solutions are well correlated with the CMT/GS-MT magnitude difference, and hence indicate magnitude uncertainty well. A normalized seismic moment tensor has four degrees of freedom and its accuracy can be represented as the non-double-couple (non-DC) component value, the 3-D angle (Φ) of DC source rotation, and a position of the rotation pole. Our results suggest that a routinely determined non-DC component is in most cases only an artifact. The distribution of the Φ-value varies over catalog time, earthquake depth, focal mechanism, and magnitude. The seismic moment errors and the value of the non-DC component are indicative of the Φ-value; for the best solutions, the 3-D angle in the CMT catalog is on the order of $5^\circ - 7^\circ$. The CMT catalog is obviously the best dataset in completeness and accuracy of its detailed solutions. Our results specifying uncertainties and completeness of global earthquake catalogs, can be used in studies of geodynamic processes, tectonic deformation associated with earthquakes, earthquake stress analysis and in many other applications of earthquake catalog data. Seismogram interpretation techniques can be reviewed and possibly revised in light of these results.

Short running title: Global Earthquake Catalogs

Key words: Earthquake catalogs, Catalog completeness, Origin time and location errors, Magnitude accuracy, Earthquake focal mechanisms and their uncertainties.

1. Introduction

Earthquake catalogs are primary data sources for inferring earthquake behavior, testing hypotheses, understanding geodynamic processes associated with earthquakes, and many other endeavors. Any quantitative study of earthquake catalogs should address the problem of catalog accuracy and its influence on obtained results. Examining catalog properties will also help guide future improvements in seismic observation and catalog reporting.

In this work we evaluate the accuracy of modern earthquake catalogs, especially the catalogs of seismic moment tensor solutions. Some catalog uncertainties like origin time, location, and magnitude have been studied earlier, primarily by comparing local with worldwide earthquake catalogs (for example, Kuge, 1992; Smith and Ekström, 1995; 1997; Röhm *et al.*, 1999; Harte and Vere-Jones, 1999; Storchak *et al.*, 2000; Patton, 2001, and references therein). For the early part of the CMT catalog Dziewonski and Woodhouse (1983a;b) reported the results of its comparison with the PDE data. Engdahl *et al.* (1998) considered global and regional hypocenter differences in the ISC and PDE catalogs and relocated nearly 100,000 hypocenters for the period 1964-1995. They also analyzed epicenter and depth shifts between the relocated solutions and that of the CMT and USGS-MT. Helffrich (1997) and Frohlich and Davis (1999) recently investigated some properties of global seismic moment catalogs.

Catalog accuracy includes a completeness of earthquake list as well as an estimate of uncertainty in determining earthquake parameters. Although some catalogs provide an internal estimate of parameter errors, such an estimate should be cross-checked by comparison with similar evaluations, preferably by independent techniques.

We use notation m (or sometimes m_w) for the moment magnitude:

$$m = \frac{2}{3} \log_{10} M - C \,, \tag{1}$$

where scalar seismic moment M is measured in Nm and $C = 6$ (*cf.* Patton, 2001). Other slightly different values for C are often used. Kanamori (1977) implies $C = 6\frac{1}{15}$, Hanks and Kanamori (1979) propose $C = 6\frac{1}{30}$, whereas, for instance, Pasyanos *et al.* (1996, p. 1264) and Ekström and Dziewonski (1988) apply $C = 6\frac{1}{15}$. (It is interesting to note, that Pasyanos *et al.* as well as Patton cite Hanks and Kanamori, 1979, as the source for their C-value). The different choice of C would only slightly shift the magnitude value by a small amount, and since magnitude is the empirical measure of earthquake size we may well select a simpler formula (1) for its computation. The magnitude is used here as an auxiliary variable mostly for illustration and comparison with old data and results. We consider the moment M-values as a primary, proper physical measure of earthquake size. In all our calculations, we take from the Harvard CMT or USGS-MT catalogs the scalar moment value M, corresponding to the double-couple source.

Quantifying errors in various catalogs is inherently difficult, since the *true* origin times, locations, and seismic moments are unknown. However, some inferences can be made. We propose to assess the differences between earthquake catalogs using three basic means. First, we compare the features of individual catalogs to certain well-verified and generally-known relations. For instance, it is generally accepted that the completeness of a catalog may be gauged by examining it in reference to the Gutenberg-Richter (G-R) law. Some

inferences about the relative completeness of different catalogs can be made by comparing the fit of each and subsets of them to such models. Another way to compare two catalogs is to examine earthquakes recorded in one catalog but not in another. We study such cases in detail and investigate the possible sources of such errors. A third method is to inspect pairs of events occurring in both catalogs. Typically an earthquake will be present in both catalogs and earthquakes common to both catalogs can be matched together. However, the matched earthquakes will differ somewhat from one catalog to the other in terms of their estimated origin time, location, scalar seismic moment or magnitude, and moment tensor. Although we cannot assume that one estimate is correct and thus measures the error in the other, some inferences may be drawn from the patterns of deviations in these estimates between catalogs.

We investigate accuracy of catalogs mostly by comparing earthquake parameters listed in different ones, as well as using the G-R law to infer a catalog completeness threshold. Catalog completeness, errors in earthquake origin/centroid time, epicentral, hypocentral and centroid coordinates, magnitude uncertainty, as well as focal mechanism and seismic moment tensor accuracy are all considered in the following sections. Since each of these earthquake parameters may become a focus for a separate extensive study, we have to limit the scope of our work. In particular, in this work we did not investigate catalog completeness in a wake of a strong earthquake (see more below in Section 3), as well as uncertainties in earthquake depth determination.

A major emphasis in our investigation is the Harvard CMT catalog, which seems best in completeness and accuracy of its solutions. The CMT catalog entries also provide the most thorough and detailed description of earthquakes, including errors in seismic moment tensor components. Of all the earthquake parameters under investigation, the seismic moment tensor uncertainties are especially targeted: The reason for this is the difficulty in studying tensor data, the analysis of which has been insufficient. Recently, regional and local MT catalogs have been compiled (see, for example, Pasyanos *et al.*, 1996; Zhu and Helmberger, 1996; Fukuyama and Dreger, 2000; Pondrelli *et al.*, 2002; Kubo *et al.*, 2002); thus accuracy analysis for MT datasets should be performed.

2. Catalogs

By modern earthquake catalogs we mean collections of estimated earthquake origin times, hypocenter or centroid locations, measures of earthquake size (scalar seismic moment or appropriate magnitude), and finally earthquake focal mechanisms or seismic moment tensors. Such data sets give a reasonably full description of an earthquake; for instance we can compute far-field, low-frequency seismic radiation or earthquake static displacement using the above information. However, detailed studies of earthquake occurrences show that this description is far from complete, since each earthquake represents a process with moment tensor or focal mechanism varying in extended time-space. Moreover, because earthquakes have fractal features, even defining an 'individual' earthquake is problematic: earthquake catalog records appear to result from a complex interaction of fault ruptures, seismographic recordings, and their interpretation (Kagan, 1991a, p. 126). We will comment on this topic later in the paper.

The catalogs satisfying the above requirements were compiled beginning in the 1970s.

Presently, several extensive catalog datasets exist. Frohlich and Davis (1999) as well as Helffrich (1997) discuss their properties. In this paper we analyze global earthquake catalogs; the companion paper (Kagan, 2002b) considers these catalogs as well as local or regional catalogs in California.

To compare catalogs of moment tensor inversions, we use two conventional (ordinary) global catalogs (PDE and ISC). The PDE worldwide catalog (Preliminary Determination of Epicenters, 1999, and references therein) is published by the USGS (U.S. Geological Survey); the current catalog ends on January 1, 2001. The catalog measures earthquake size, using several magnitude scales, of which the body-wave (m_b) and surface-wave (M_S) magnitudes are provided for most moderate and large events since 1965 and 1968, respectively. The catalog contains more than 50,000 shallow earthquakes with $m_b \geq 5$ from 1965 to 2001.

We also analyze the earthquake distributions for the ISC worldwide catalog (International Seismological Centre, 1995, and references therein). The ISC catalog supplies solutions from many seismographic networks. In this study we use only the ISC solution for all earthquakes, the records identified as explosions are excluded. The available catalog starts 1964/1/1 and ends 1998/12/31.

We study the earthquake distributions for the global catalog of moment tensor inversions compiled by the Harvard group (Dziewonski *et al.*, 2001; see also references and earthquake statistics for 1977-1998 in Dziewonski *et al.*, 1999). Although the current catalog starts from 1976 (Ekström and Nettles, 1997), the 1976 data are significantly different from the rest of the catalog in terms of their completeness, thus in this work we use the catalog starting from 1977. The catalog contains 17,665 solutions over a period from 1977/1/1 to 2000/12/31. In addition to origin times, hypocentral coordinates and magnitudes m_b and M_S taken from the PDE or ISC catalogs, the CMT catalog includes seismic moment centroid times and locations as well as estimates of seismic moment tensor components (Dziewonski *et al.*, 1981; Dziewonski and Woodhouse, 1983a;b). Each tensor is constrained to have zero trace (first invariant), i.e., no isotropic component. Double-couple (DC) solutions, i.e., with tensor determinant equal to zero, are supplied as well. Almost all earthquake parameters are accompanied by internal estimates of error.

The global catalog of moment tensor solutions, issued by the USGS (Sipkin, 1986; Sipkin *et al.*, 2002, and references therein), spans 1980/1/1–2000/7/1, and contains 2923 solutions. We call this catalog GS-MT. The origin times and hypocentral coordinates are taken from the PDE catalog, and the seismic moment tensors are assumed to have zero first invariant. The catalog does not provide estimates of tensor component errors.

The global catalog of first-motion focal mechanisms was also issued by the USGS (Needham, 1988, and references therein). The available catalog, called GS-FM below, spans 1981/1/1–1990/5/1 and contains 983 DC solutions.

3. Catalog completeness

Catalog completeness needs to be considered from three somewhat interconnected points of view:
(1) Due to limited sensitivity and coverage of the Earth by seismographic networks, small events are generally missing from earthquake catalogs. Thus, we need to establish a mag-

nitude threshold (or cutoff) m_t, above which a catalog can be considered reasonably complete. We use the deviation from the G-R relation to study catalog completeness (Section 3.1).
(2) The magnitude threshold is not uniform in time and space; the largest level fluctuations occur after large earthquakes. For example, a cursory inspection of the southern California catalog (Hileman *et al.*, 1973) demonstrates that after the 1952 Kern County $M_L = 7.7$ earthquake, the threshold increased to about $m_t = 4.0 - 4.5$ from its usual value $m_t = 3.0$ for at least a few days. In active aftershock sequences after a strong earthquake, an inter-event interval is often smaller than an earthquake duration, hence earthquake records overlap, presenting a difficult problem for event identification and interpretation. With a high probability an overlapping seismogram record would be identified as a larger shock, leaving a catalog depleted of smaller aftershocks. In addition, catalog compilers are overwhelmed by a great number of aftershocks in a sequence, so they may pass over some smaller shocks. The global catalogs we study in this work have a much smaller magnitude range, and are thus less vulnerable to such disruptions. In principle, magnitude threshold fluctuations can be studied by the same methods we use in item 1. We do not investigate these intermediate-term (hours-days-weeks) threshold fluctuations here.
(3) Earthquakes are not point sources, but have complex temporal and spatial internal structure. Therefore, depending on the frequency of recording equipment and interpretation method, each can be represented as one earthquake or as an extensive foreshock/mainshock/aftershock sequence (Kagan, 1991a). These short-term (seconds-minutes) differences in the catalog completeness are considered in Section 3.2; we use the comparison of missing earthquakes in different datasets to study these catalog features.

3.1 Magnitude completeness threshold

Several methods have been proposed to establish the magnitude completeness threshold for local and global catalogs. Wiemer and Wyss (2000) review many such methods and propose their own technique based on fitting the magnitude distribution with an exponential law (an analog of the G-R relation) and estimating the average difference between a theoretical approximation and experimental statistical distribution. The G-R distribution is equivalent to a power-law or the Pareto law for scalar seismic moment values (Kagan, 1997; 2002a).

In this work we first apply a simple method for testing a possible non-uniformity of the GS-MT (and other) catalogs: we compare the numbers of shallow (depth limits 0–70 km) events in about 1 year time intervals (Fig. 1) throughout the catalog's time span (see also Fig. 1 in Kagan, 1997). Such a test should yield a reasonable result if there are few extended aftershock sequences in a catalog: a likely condition for a catalog with a high magnitude cutoff like the Harvard and GS-MT datasets. The curves in Fig. 1, especially for smaller magnitudes, exhibit a highly inhomogeneous earthquake rate; only after 1995 do the annual numbers of earthquakes stabilize for $m \geq 5.6$ events. Only for $m \geq 6.2$ is the rate stable starting at about 1983, i.e., during most of the GS-MT catalog's time-span.

For the Harvard catalog we employ a technique similar to Wiemer and Wyss' (2000) design, but based on a more conventional statistical method. It estimates not only the difference between theoretical and observed distributions, as in the above reference, but

also a probability that this difference may be due to random fluctuations. We approximate the seismic moment distribution (Kagan, 2002a) by a power-law (in agreement with the G-R relation) when changing the magnitude threshold from some minimum value up to $m_v = 6.8$, i.e., clearly above any possible value of the catalog's lower threshold. Therefore, this method assumes that 1) earthquake size has a theoretical power-law distribution and 2) deviation from this law at the small earthquake side of the distribution is caused by some events missing in a dataset.

In Fig. 2 we display the magnitude-frequency relation for earthquakes in three depth ranges: shallow, intermediate, and deep. Approximation by a G-R law is also shown. The curves (solid lines) exhibit random fluctuations, especially at the large earthquake end.

We use these distributions to determine the sliding value of β (the power-law exponent) as the lower magnitude cutoff moves from $m_v = 4.7$ (the minimum moment magnitude recorded in the catalog). Fig. 3 shows β depending on m_v for three depth ranges in the CMT catalog. The curves behave similarly, starting with a very low value at $m_v = 4.7$. Around $m_v = 5.5$, the β-values stabilize at the $\beta \approx 2/3$, which corresponds to the classical b-value of the G-R law equal to 1.0. After the plateau, β starts increasing again around $m_v = 6.2$. This gradual increase is due to the simple power-law's inability to account for decay in the distribution tail more quickly than a power-law. This decay is thought to come from the finite size of the earthquake generating tectonic layers (see more in Kagan, 1997; 2002a).

At the smallest values of m_v, the power-law (or G-R relation) does not fit the empirical distribution. To account for this, we employ the Kolmogorov test (see, for example, Stephens, 1974; Kagan, 1995) assuming that the seismic moment empirical distribution follows the power-law with the exponent β (the null-hypothesis). The test measures the probability α of making an error in rejecting the hypothesis. For example, $\alpha = 0.01$ means that we may reject the null-hypothesis when in reality it is true, but the probability of such an error is low – 1%.

There is a problem with the standard Kolmogorov test: the null-hypothesis should be known exactly. If we determine the parameter β from experimental data, the test needs to be modified (Stephens, 1974). Kagan (2002a) suggests that for all earthquakes the parameter β has a universal value of about 0.63. If we assume this conjecture, we can employ the standard Kolmogorov test. However, even if the earthquake size distribution follows a universal law, Kagan (2002a) shows that the maximum earthquake size or earthquake corner seismic moment (M_c) varies, especially for oceanic earthquakes. If several earthquake populations with the universal β-value but with different M_c-values are mixed, the resulting dataset would exhibit apparent β variation (Vere-Jones *et al.*, 2001).

Below we carry out the Kolmogorov test using both assumptions: the universal β-value and the β-value determined using an equivalent of Aki's (1965) formula (Deemer and Votaw, 1955). To apply the Kolmogorov test in practical terms, we arrange our earthquake size data in ascending order and then calculate for the k-th earthquake

$$D_N^- = \max_{1 \le k \le N} \left(\frac{k}{N} - F(M_k) \right) , \qquad (2)$$

which is the maximum difference between the theoretical $F(M_k)$ and empirical cumulative

distribution functions. Here N is the number of earthquakes in a catalog. Although there could be an 'excess' of earthquakes near the threshold value due to chance, we assume that such an event has a negligible probability of reaching a critical level and can be ignored. Thus, because missing earthquakes would cause the theoretical curve to fall below the empirical distribution function near the magnitude threshold, we need calculate only a one-sided difference.

Standard statistical tables provide α-values for selected D_N^- differences. In our test we want to calculate the variation of α as the magnitude threshold varies continuously. Bol'shev (1963, p. 143) proposes for $0.2 > \alpha > 0.01$ and $N > 50$ the following approximation for the Kolmogorov distribution:

$$D_N^- \simeq \sqrt{\frac{-\log \alpha}{2N}} - \frac{1}{6N}\,, \tag{3}$$

From (2) and (3) the significance level α can be calculated as

$$\alpha \simeq \exp\left[-2N\left(D_N^- + \frac{1}{6N}\right)^2\right]\,. \tag{4}$$

As outlined above, these equations are valid for the standard Kolmogorov test when the theoretical distribution is known exactly. Table 1A in Stephens (1974) shows that for the modified Kolmogorov test, in the range $[0.2 \geq \alpha > 0.01]$ the D_N^--values in (4) should be increased by a factor of about 1.2 to account for estimating the β-parameter.

In Fig. 4 we display the α-values for earthquakes in three depth ranges. The behavior of curves is similar for this and other plots: usually the α-level raises rapidly at some m_t-value which we call the 'critical' or 'real' threshold value. Before this sharp increase, the null-hypothesis can confidently be rejected, but if we reject it afterwards for larger m_v, we risk making a large error: rejecting the hypothesis when in reality the power-law model is true. For larger m_v-values sometimes the α-curves decay again (as in Fig. 4 for intermediate and deep events). This decay is not caused by missing earthquakes, because if a catalog is reasonably complete at a threshold m_t-value, it is unlikely that even larger events have been missed. The cause of the α-value decay may be random fluctuations, some details of seismogram interpretation technique, or perhaps the inhomogeneity of earthquake populations discussed earlier in this subsection.

For example, the Harvard catalog uses long-period 'mantle' waves to interpret larger earthquakes (Dziewonski *et al.*, 1981; Dziewonski and Woodhouse, 1983a;b; Kuge and Lay, 1994a; Frohlich and Davis, 1999). Depending on its particulars, this technique may introduce a bias in determining the scalar seismic moment, and subsequently a deviation from a power-law behavior around a magnitude where such a change is taking place. Earthquakes included in the GS-MT catalog are also processed differently depending on their size (Sipkin, 1994; Frohlich and Davis, 1999); see more discussion in Section 6.1. Similar influences of seismogram processing technique on magnitude distribution are unavoidable, and special care must be taken to minimize and correct them if necessary.

To demonstrate the possibility of magnitude determination bias, Fig. 2 displays magnitude-frequency relation for earthquakes in three depth ranges: shallow, intermediate, and deep. One set of curves is for all earthquakes in the CMT catalog; another is for

earthquakes in which mantle waves have been used in solutions. These low-frequency (period 135 s) waves are used mostly for larger earthquakes. From the diagram, the transition from employing purely body-wave solutions to both type of waves occurs at about $m = 6.3$ for shallow events, $m = 6.5$ for intermediate ones, and $m = 6.7$ for deep earthquakes.

Tables 1 and 2 display the magnitude threshold values obtained for various time, depth, Flinn-Engdahl region categories (Kagan, 1997), and focal mechanisms of earthquakes. From Table 1 it is obvious that both Kolmogorov tests described above produce similar results. The threshold is higher in the first five years of the Harvard catalog and decreases afterwards, confirming the results of more informal analysis (Molchan *et al.*, 1997; Kagan, 1997; 2002a). The annual numbers of events shown, for instance, in Dziewonski *et al.* (1999, their Fig. 1) also confirm this pattern: the numbers are significantly lower for 1977-1981, increase for the 1982-1986 period and stabilize thereafter. Intermediate and deep earthquakes have in general a lower value of m_t. Since the accuracy of the test depends on the number of earthquakes sampled, estimating the magnitude threshold for these events exhibits higher random fluctuations.

Table 2 shows the m_t-value estimates for earthquakes in different tectonic provinces and having different focal mechanisms. For instance, we combine into 'Subduction' region category earthquakes occurring in several Flinn-Engdahl regions with subduction type plate boundary deformation (see more detail in Kagan, 1997 and 1999).

An earthquake is considered to have a normal focal mechanism if its most-compressive principal axis of the moment tensor (P-axis) is more vertical than either principal axis B or T (Frohlich 2001). Similarly, we determine focal mechanisms for thrust and strike-slip events. Again, since the numbers of earthquakes in each table entry is relatively small – usually a few tens or hundreds of events (Table 5 in Kagan, 2002a) – the m_t-values exhibit great random variability. Subduction and oceanic earthquakes have relatively large m_t-values, whereas continental events have the smallest magnitude threshold. Except for strike-slip oceanic earthquakes, there is little obvious difference between earthquakes of various focal mechanisms. The high m_t-value for the strike-slip oceanic earthquakes can be explained by the low and varying corner moments of these events (Bird *et al.*, 2002). Testing this strongly inhomogeneous sample yields a high magnitude cutoff estimate, since the curvature of the magnitude-frequency relation caused by varying corner magnitude m_c resembles that produced by incompleteness of the catalog. When evaluating a catalog's completeness by the Kolmogorov test we assume that earthquakes follow the G-R relation; if the corner moment is close to the threshold value, there is practically no scale-invariant region in earthquake size distribution, hence the test results are biased. Thus, the high m_t-value may be an artifact here.

3.2 Catalog comparison and temporal features of earthquake occurrence

As another test of catalog completeness, we employed a program in two catalogs which matched events closest in time and space. After some experimentation the following windows were selected for comparing the CMT and GS-MT catalogs: a distance between centroid and hypocenter of less than 140 km, a time difference of less than 1 min. The GS-MT catalog lacks many earthquakes from the CMT catalog, especially before 1995,

Fig. 1 explains this by a high magnitude cutoff for the former catalog before 1995. Thus, below we primarily investigate the completeness of the CMT catalog.

Fig. 5 displays a distribution of earthquake magnitudes for the CMT catalog which cannot be matched with the GS-MT events using the criteria outlined above. The lack of earthquakes with $m < 6$ in the CMT dataset could be explained by closeness of these earthquake magnitudes to a magnitude threshold and magnitude uncertainties in both catalogs (see Section 6 below): while the GS-MT magnitude may be above, for instance 5.8, the CMT magnitude may be lower, putting it below m_t.

We inspected individually all 'missing' earthquakes $m \geq 6$ in the plot. Out of seven events, the most plausible explanation for six mismatches is that according to the PDE catalog, these earthquakes have a complex temporal structure: two (or several) events are listed in the PDE catalog for each CMT entry. In some cases the GS-MT catalog selected a close foreshock as a mainshock, whereas the CMT catalog would use a second earthquake in a sequence as the main event, or *vice versa*. Only one earthquake, 1994/7/25 22:0:23, $m = 6.62$ listed in the GS-MT catalog appears to be totally missing from the CMT catalog.

Identifying earthquakes becomes more difficult when we compare the CMT catalog with the PDE or ISC datasets. Christophersen (2000, her Fig. 2.8) indicated that even some of PDE and ISC $m \geq 7$ earthquakes are 'missing' in the Harvard catalog. Since the PDE and ISC catalogs are based on higher frequency seismic recording, they often list several earthquakes for each CMT event. Most of these could be classified as close aftershocks and sometimes as close foreshocks. We found 50 earthquakes $M_S \geq 6$ present in the PDE catalog but missing in the CMT dataset which can be explained as close aftershocks. Some of these earthquakes can be quite large: seven events have $M_S \geq 6.5$. The largest of these events, the New Ireland region earthquake 2000/11/16 at 7:45:32, recorded $M_S = 7.2$ in the PDE catalog. This earthquake is clearly a close aftershock of an event at 7:42:17, $m = 7.87$, which is listed in the CMT dataset.

About 30 $M_S \geq 6$ events in the PDE catalog cannot be explained by temporal proximity to CMT earthquakes. We investigated individually all earthquakes with $M_S \geq 6.3$ in the PDE catalog missing from the CMT catalog. Upon close inspection, most appear to be part of foreshock-mainshock-aftershock sequences. Only two moderate earthquakes – 1977/11/24 2:2:31, $M_S = 6.3$ and 1985/3/4 3:32:49, $M_S = 6.6$ are missing entirely from the CMT catalog. The latter event is obviously an aftershock of a strong Chilean earthquake 1985/3/3 22:47:06 ($m = 8.0$).

We carried out a similar comparison of the CMT and 1977-1998 ISC catalogs. There are 45 earthquakes with $m_b \geq 6$ in the ISC catalog which cannot be matched to CMT events. Almost all of these can be identified with close aftershocks or foreshocks of CMT events. Two earthquakes (1978/5/5 23:22:36, $m_b = 6.3$ and 1994/7/25 22:0:45, $m_b = 6.3$) lack appropriate equivalents in the CMT catalog, and the first event is missing in the PDE list as well. The second earthquake was mentioned above as listed in the GS-MT catalog and missing in the CMT set.

We investigate quantitatively the difference in earthquake time distribution for the global catalogs. Fig. 6 displays an example of a distribution of inter-earthquake time intervals for the CMT catalog. The cumulative distribution is binned according to the magnitude of the earthquake that was first in time; for larger events, longer time intervals

are observed. This feature may be explained by longer coda waves for large earthquakes. The time delay between earthquakes could be characterized as 'dead-time'. Of course for earthquakes such dead-time is not a deterministic, clear-cut feature, it manifests itself by statistical shift of a distribution toward longer time intervals. The lack of sharp cutoff in distribution plots can probably be explained by occurrence of strong aftershocks in a coda of mainshocks, or even an earthquake larger than an initial shock. In such cases, seismogram interpreters would try to identify and process a large event, an effort less likely for a smaller aftershock (see also discussion in item (2) in the beginning of this section).

We summarize the results of measuring inter-earthquake time distributions in Table 3. The decile $\tau_{0.1}$ is the time interval in which 10% of earthquakes follow another earthquake in $6.0 > m \geq 5.5$ magnitude range. Such an interval approximately characterizes the time delay caused by coda waves from the first event on identifying and listing the following earthquake. The τ-parameter value is dependent on frequency of waves used in magnitude determinations, 1 sec for m_b, 20 sec for M_S and several tens or hundreds of seconds for moment magnitude. These greatly different values are at least partly responsible for identification of earthquakes in close sequences of events. Kagan (1991a, p. 126) discusses the 1971 San Fernando, California earthquake for which widely different aftershock numbers are available in various catalogs.

As another example, let us consider aftershock numbers for the 1999 Chi-Chi (Taiwan) earthquake: the local catalog (Teng *et al.*, 2001, p. 902) lists 17 $M_L \geq 5.0$ events in the first hour of the sequence, some of them quite large, including four with $M_L \geq 6.4$. For the same time period, the PDE catalog shows six $m_b \geq 4.5$ aftershocks. Finally, the first aftershock in the CMT dataset comes only four hours after the mainshock. Therefore, depending on frequency characteristics of seismographic network and seismogram processing technique, the same earthquake sequence may be identified as one complex earthquake with some subevents, or as a foreshock/mainshock/aftershock sequence with many 'individual' earthquakes.

The different treatment in various catalogs of close-in-time foreshocks and aftershocks is especially important in studying earthquake temporal relations. During an earthquake and immediately before and after it, the rate of dependent events increases as a power-law according to Omori's law (Kagan, 1991b; Ogata, 1998; 1999). During this period any 'missing' earthquake would strongly influence an estimate of occurrence rate, so proper accounting and modelling of these events is a necessary pre-condition for any such investigations. To ensure catalog completeness and homogeneity, in earlier investigations (Kagan, 1991b, p. 145) we excluded from a catalog every aftershock which is closer to a mainshock than the estimated coda wave duration. Unfortunately, this important property of earthquake catalogs has not been sufficiently studied.

Summarizing the results of a completeness study for the Harvard catalog, we infer from our tests that the catalog appears to be reasonably complete. If we omit earthquakes in temporarily close foreshock/mainshock/aftershock sequences which are not resolved by low frequency seismograms used in the CMT processing technique, only a few moderate ($6.7 > m > 6.2$) earthquakes may be missing from the catalog. It seems that these events occurred primarily in the early part of the catalog. The magnitude completeness threshold of the catalog decreased from about 5.7 in the first 5 years to about 5.4 in recent years. The

threshold also seems to vary with earthquake depth, tectonic region, and focal mechanisms of earthquakes.

4. Origin/centroid time

Fig. 7 displays time difference ($\Delta t = t_{\rm cmt} - t_{\rm pde}$) for CMT/PDE shallow (0-70 km) matched earthquakes *versus* moment magnitude. We approximate the distribution by a linear and quadratic relation Δt *versus* magnitude. One should expect that Δt is proportional to $M^{1/3}$ (Smith and Ekström, 1995), but since some Δt-values are negative (i.e., centroid time is less than the origin time), we cannot use the power-law in the approximation. Instead, we binned Δt into half-unit magnitude intervals and, if the earthquake number in a bin exceeds four, we show in Fig. 8 the dependence of average Δt and its standard deviation on magnitude. By a solid line we display a power-law approximation obtained by Smith and Ekström (1995) when comparing the CMT and ISC catalogs $\left[\Delta t = 3.0 \times 10^{(m-6)/2)}\right]$. The above formula is similar to that obtained by Dziewonski and Woodhouse (1983b, their Figure 13a) for shallow earthquakes.

There is a 2 s baseline shift in theoretical teleseismic P-wave travel times between the PREM model (Dziewonski and Anderson, 1991) used by the Harvard group and the Jeffreys-Bullen model used by the ISC and PDE (Smith and Ekström, 1995; Ekström, private communication, 2002). Following Smith and Ekström (1995), we adjust the time Δt by 2 s to account for this bias.

The logarithmic scale of Fig. 8 misrepresents the relation between average time $<\Delta t>$ and its standard deviation $\sigma_{\Delta t}$. In reality the coefficient of variation $\Delta t/\sigma_{\Delta t}$ is approximately constant (about 0.4-0.5) for all magnitude bins. Such a relation is to be expected if earthquake rupture is self-similar in time-space (*cf.* Smith and Ekström, 1997, their Fig. 1).

In Table 4 we display time difference (Δt) for CMT/PDE and PDE/ISC catalogs as approximated by a linear relation Δt *versus* magnitude. Results from Fig. 7 are entered as the second row in the table. Most earthquakes in Fig. 7 are concentrated in the magnitude interval $6 - 7$, where the distribution exhibits no effects of non-linearity (i.e., linear approximation is like the quadratic one). The average time delay (Δt) for $m = 6$ event is about 4 sec; the standard deviation decreases slightly with depth. There is no appreciable difference in the Δt behavior for CMT/PDE or CMT/ISC catalog comparison.

As should be expected, comparing the PDE and ISC catalogs yields different results: the average Δt is close to zero, and its standard deviation is smaller than that of the CMT/PDE and CMT/ISC match. Both Δt and $\sigma_{\Delta t}$ for PDE/ISC comparison show practically no dependence on magnitude. In the CMT/PDE and CMT/ISC comparisons, Δt includes both time errors and the effects of rupture propagation from a hypocenter to a centroid. For the PDE/ISC comparison, only errors in origin time contribute to Δt; however, since both these catalogs often use data from the same seismographic stations, the $\sigma_{\Delta t}$ value may be an underestimate. The standard error again decreases with earthquake depth. To estimate origin time errors (σ_t) in the PDE and ISC catalogs, we may assume that such errors are approximately equal in both datasets. Then

$$\sigma_t \;=\; \sigma_{\Delta t}/\sqrt{2}\,, \tag{5}$$

This implies that the average origin time uncertainties are on the order of one second for

the PDE and ISC catalogs.

There is little, if any, change in the CMT/PDE or PDE/ISC time difference (Δt) during 1977-2000 and 1971-1998 time span of both catalogs, respectively. However, the standard error ($\sigma_{\Delta t}$) decreases slightly – by about 10% – during 1971-1998 for PDE/ISC, possibly reflecting more accurate solutions.

5. Location uncertainty

We compare earthquake locations in three catalogs: CMT, PDE, and ISC by measuring horizontal distances between the matched events, i.e., we study relative location differences. For the CMT and PDE catalogs, Fig. 9 displays the dependence of average distances between a horizontal projection of the centroid and the epicenter on the moment magnitude. As in Fig. 8 we exclude the last bin that contains only three earthquakes (see Fig. 7). Although the distribution of distances cannot be Gaussian, we also show standard deviations to display the uncertainty of each distance measurement. The standard errors and average distances decrease slightly with magnitude in the magnitude range $[6-7]$; afterwards they start to increase. The decrease is apparently caused by a better location accuracy for stronger earthquakes, whereas a subsequent increase is connected with a finite size of earthquake focal zone and difference in centroid and epicenter locations (Smith and Ekström, 1997). At magnitude about $7-7.5$ the latter effect is stronger, so the last few points are better approximated by a power-law line suggested by Smith and Ekström (1997) for dip-slip earthquakes ($L = aM^{1/3}$, with $a = 2.79 \times 10^{-8}$, where moment is measured in *dyne-cm*).

For smaller earthquakes, a simple quadratic polynomial formula yields a better fit in Fig. 9. A minimum value for average distance (about 30 km) is reached at about magnitude 7, consistent with findings by Smith and Ekström (1997, their Fig. 2). They state that the minimum median distance for CMT and ISC locations is approximately 25 km. Our comparison of CMT and ISC catalogs yielded distance values slightly (5-10%) higher that those for the CMT/PDE match.

Average CMT/PDE distances decrease with earthquake depth: for $m = 6$ and depth interval 0-70 km, we obtain average distance $< R > =$ 37.4 km; for depth interval 70-300 km $< R > =$ 31.4 km; and for depth interval 300-700 km $< R > =$ 26.7 km. Engdahl *et al.* (1998, their Table 2) also find that the average epicenter difference between the CMT solutions and their relocated epicenters decreases with depth from about 34 km for shallow earthquakes to about 31 km for deeper events.

The distance significantly depends on catalog earthquake time T in years,

$$< R > = \; a_0 + a_1 \times (T - T_c)\,, \tag{6}$$

where $a_0 = 39.1$ km, $a_1 = 0.44$ km/y, and $T_c = 1977$. These $< R >$ regularities mean that as of now the distance between centroid and epicenter should be of the order 20-25 km for $m = 7$ shallow earthquakes and decrease to about 20 km for deeper events.

Fig. 10 displays the average distances for matched epicenters of the PDE and ISC catalogs. As expected, there is little or no dependence of $< R >$ on magnitude for the largest earthquakes, and the average distances and standard deviations decrease in magnitude interval $m_b = 5-6$, again due to more accurate earthquake locations. For $m = 6$ shallow

earthquakes the average distance is about 7 km, testifying to reasonably accurate epicenter locations in these catalogs. One must note, however, that earthquake data in these catalogs are not fully independent. Both catalogs use, at least partly, the same set of seismographic records and the same or similar Earth structure in epicenter construction (Storchak *et al.*, 2000).

Similar to the CMT/PDE match, the average PDE/ISC distances also decrease with earthquake depth and exhibit a similar dependence on time: for $m = 5$ and depth interval 0-70 km we obtain $a_0 = 13.2$ km (see Eq. 6, these a_0-values are for year $T_c = 1971$); for depth interval 70-300 km $a_0 = 10.9$ km; and for depth interval 300-700 km $a_0 = 11.1$ km. The value of a_1 is almost the same for all depth ranges $a_1 = 0.22$ km/y; hence, present-day accuracy should be significantly better than the above values.

Fig. 11 displays the histogram of epicentral distances for PDE/ISC matched events. If errors in epicenter location have a Gaussian distribution with a standard error σ_x and σ_y in latitude and longitude, respectively, the distances should be distributed according to the Rayleigh distribution:

$$\psi(R) = \frac{R}{\sigma_R^2} \times \exp\left[-R^2/(2\sigma_R^2)\right] \ , \tag{7}$$

where σ_R is an uncertainty in latitude or longitude difference for epicentral distances: $\sigma_R = \sigma_x\sqrt{2}$, if errors are equal for the PDE and ISC catalogs, and $\sigma_x = \sigma_y$. The Rayleigh law corresponds to the distribution of a vector length in two dimensions, if the components of a vector have a Gaussian distribution with zero mean and the standard error σ_R. The theoretical curve in Fig. 11 has been adjusted to have the same average distance $< R >$ as in the empirical distribution.

The fit of the theoretical curve to the histogram is far from good; the empirical distribution has a much 'fatter' tail than the Rayleigh law. Why is this?
(1) Some events may be misidentified in two catalogs or the epicenter coordinates may have been determined with a large error in one of the datasets (*cf.* Storchak *et al.*, 2000). Individual manual inspection may be necessary to review all these outliers and correct them, if necessary (*cf.* Smith and Ekström, 1995; 1997). Because of the large volume of work required, we have not yet resolved this problem.
(2) The fatter tail may result from location errors that are strongly inhomogeneous in time and space (Storchak *et al.*, 2000). The evidence outlined above indicates that the average distance varies in time and depends on earthquake depth. It should be expected that for some earthquake pairs, such distance variations may be greater than others and so cause the distribution tail to 'fatten' up.

6. Magnitude accuracy

Before analyzing magnitude errors, we should discuss defining magnitude as a measure of earthquake size. The seismic moment and its derivative, moment magnitude (see Eq. 1), have a clear physical meaning (Aki and Richards, 1980), which makes them appropriate to measure earthquake size. However, as we mentioned in Section 3, the fractal nature of earthquake process makes defining an individual earthquake and its parameters uncertain for close-in-time sequences.

Ordinary magnitudes are empirical functions developed to represent earthquake size in the early days of seismology. These quantities lack a well-defined physical meaning: one usually uses empirically derived correlations to obtain physical quantities from conventional magnitudes. Thus, in evaluating magnitude errors, two factors need to be gauged: the uncertainty of a magnitude measurement itself and any error in transforming the obtained magnitude into a physically relevant quantity.

6.1 Moment tensor catalogs

Fig. 12 displays the dependence of difference in moment magnitudes for matched shallow earthquakes in the CMT and GS-MT catalogs. The dependence has been approximated by a linear

$$\Delta m \;=\; a_0 + a_1 \times (m - m_r)\,, \tag{8}$$

and quadratic regression

$$\Delta m \;=\; a_0 + a_1 \times (m - m_r) + a_2 \times (m - m_r)^2\,, \tag{9}$$

where $m_r = 6$, and m is the moment magnitude from the CMT catalog ($m \geq 6.0$). If magnitudes differ by only random error, the values of regression coefficients a_0, a_1, and a_2 should be close to zero. Their values are small, only for the quadratic polynomial a_2 is relatively large, pointing out the non-linearity of the mutual magnitude relation. Sipkin (1994, p. 1667) notes that the filter employed by the USGS inversion technique "... can lead to a bias in the estimated scalar moments larger than approximately 10^{20} Nm [$m > 7.33$]." Most likely, this is the bias observed in the plot. Another source of the non-linearity is a change in the magnitude evaluation technique as exemplified in Fig. 2. The variance for the magnitude difference is the result of errors in both catalogs,

$$\sigma^2_{\Delta m} \;=\; \sigma^2_{\rm cmt} + \sigma^2_{\rm usgs}\,, \tag{10}$$

where $\sigma_{\rm cmt}$ is the assumed standard variation for the CMT catalog.

In Table 5 we collect the values of the regression coefficients and standard errors for several choices of catalog, depth ranges, and magnitude comparison. Results from Fig. 12 are entered as the first row in the table. Table 9 in Kagan (2002a) lists CMT/GS-MT log-moment differences ($\Delta \log_{10} M$) for other magnitude thresholds and catalog time spans.

A similar plot (Fig. 13) for the CMT/PDE catalogs shows a different picture: all the regression coefficients are relatively large, testifying to the large difference in both magnitudes (compare also Table 5, row 10). The standard error $\sigma_{\Delta m}$ is also by almost a factor of two larger than for the CMT/GS-MT comparison. The significant non-linearity of the quadratic regression curve comes from the saturation of the M_S scale, which starts at about $M_S = 7.5$ (*cf.* Ekström and Dziewonski, 1988).

For the body-wave magnitude m_b, the coefficient values and the standard error increase again (Table 5, row 6). Fig. 14 displays the comparison in the more conventional format m_b *versus* m. The m_b saturation effect is also represented more clearly than in Fig. 13. The line of dots starting, for example, at $m_b = 6$ shows that a $m_b = 6$ earthquake may correspond to the moment magnitude 7.5 and larger: the body-wave magnitude is of little value in estimating seismic moment for a large earthquake (*cf.* Dziewonski and Woodhouse, 1983b).

Since the $\sigma_{\Delta m}$-value for CMT/PDE comparison is significantly larger than that for CMT/GS-MT, it is clear that most of the uncertainty is caused by m_b errors, conversion errors $m_b \rightarrow m$, or combination of both. If we take, for example, $\sigma_{m_b} = 0.25$, then the uncertainty in estimating the moment magnitude from known body-wave magnitude is about

$$\tilde{\sigma}_m \approx \sigma_{m_b}/a_1 = 0.41\,, \tag{11}$$

assuming a linear fit. For the quadratic regression the slope of the curve is even smaller for $m > 7.5$; hence, the errors would be significantly larger than (11) predicts.

Table 5 summarizes the main results of magnitude accuracy analysis. The first four rows demonstrate that the magnitude uncertainty decreases by about a factor of two as earthquake depth increases from about zero to 700 km. In all comparisons of moment magnitudes (CMT/GS-MT) the bias (a_0) and linear term (a_1) values are very small, demonstrating that for moderate and large earthquakes, moment tensor inversion provides a consistent measure of earthquake size. For only the largest earthquakes, as relatively large values for the quadratic regression term (a_2) demonstrate, the GS-MT moment magnitude apparently underestimates earthquake size. The comparison results for 1995-2000 (row 5) indicate that after 1995 moment magnitudes in the GS-MT catalog are much closer to the CMT values, especially for large earthquakes.

Magnitude comparison of MT and conventional catalogs (PDE and ISC) – rows 6-9 and 15-18 – again demonstrates the higher accuracy of magnitude determination for deeper events: the $\sigma_{\Delta m}$-values decrease with depth. Large values for all regression coefficients testify to a significant bias in m_b and M_S determination, as well as saturation effect for both these magnitudes. If one knows the bias, linear and quadratic term values (i.e., a_0, a_1, and a_2), one can correct these magnitudes to obtain the moment magnitude estimate. But the discussion around (11) indicates that the obtained moment magnitude values would be highly inaccurate, especially for m_b–m_w conversion.

To evaluate magnitude accuracy of both MT catalogs, we compared them with a third catalog, PDE or ISC. The values of $\sigma_{\Delta m}$ for CMT/PDE and GS-MT/PDE are comparable (rows 11-14 in Table 5). We used the time period 1995-2000 for our test, since the GS-MT catalog seems to have a lower and stable magnitude threshold during this time (Fig. 1). Similar results are obtained for a CMT/ISC and GS-MT/ISC comparison (rows 15-18 in Table 5). The approximate equality of the $\sigma_{\Delta m}$-values suggests that magnitude errors for both MT catalogs are equal, so we can estimate an error for an individual magnitude measurement (σ_m), using a formula similar to (5).

$$\sigma_m = \sigma_{\Delta m}/\sqrt{2}\,. \tag{12}$$

The magnitude accuracy of both MT catalogs appears to be on the order of 0.09 for shallow earthquakes, decreasing to about 0.05 for deep events. These values generally agree with differences in log seismic moment for the CMT and GS-MT catalogs, reported by Helffrich (1997). He quotes the following values for the standard deviation: 0.15, 0.06, and 0.07 for shallow, intermediate and deep seismicity, respectively. If we divide the above values by 1.5 (see Eq. 1), we find that Helffrich' estimates for deeper events are slightly smaller than ours.

We estimate the dependence of the CMT/GS-MT moment magnitude difference and its standard deviation on magnitude and catalog time. The standard deviation $\sigma_{\Delta m}$ for all earthquake depths (0-700 km) increases with the magnitude from values of 0.1 for $m = 6$ to about 0.15 for $m = 8$. This increase is, most probably, connected to a bias of the GS-MT catalog in estimating seismic moment for very large earthquakes (see above).

The variations in Δm and $\sigma_{\Delta m}$ over time are displayed in Fig. 15. From the plot it is clear that whereas the average magnitude difference exhibits no clear pattern, the standard deviation decreased significantly during the time-span of both catalogs. Starting from $\sigma_{\Delta m}$ about 0.15 in 1980-82, the standard deviation reached about 0.08 in the most recent five years.

6.2 Conventional earthquake catalogs

Comparing the PDE and ISC catalogs (rows 19-21 in Table 5) shows that m_b estimates are reasonably close: the standard error for the difference is similar to that for MT catalogs (see also Christophersen, 2000, section 2.5.2). This suggests that the uncertainties in m_b determination are relatively low, i.e., on the order of 0.09 magnitude units (see Eq. 12). However, the difference between two standard magnitudes, m_b and M_S, is large; standard deviations are comparable to the $\sigma_{\Delta m}$-value for comparison with moment magnitudes.

By comparing the $\sigma_{\Delta m}$-values for different pairs of magnitudes, we can roughly estimate magnitude errors separately. We discussed the moment magnitude errors above; these estimates are significantly smaller than $\sigma_{\Delta m}$-values for m/m_b and m/M_S comparison. Therefore, we assume that m_b and M_S conversion errors contribute mainly to $\sigma_{\Delta m}$. Hence, using the expression similar to (10), we estimate the m_b total error as about 0.25 and M_S standard error about 0.2. Ekström and Dziewonski (1988) obtained a similar uncertainty estimate for $M_S - m$ conversion.

These standard variation estimates assume that errors are random, i.e., there is no systematic effect in magnitude determination. These assumptions are clearly not valid: catalog compilers use largely overlapping sets of seismograms, so the standard error may underestimate real magnitude uncertainty. For example, the difference between m_b-values in the PDE and ISC catalogs is relatively small, but as we see from other comparisons, the m_b-magnitude imprecisely measures an earthquake's size as indicated by its seismic moment. Ekström and Dziewonski (1988) remark that the M_S values have significant regional variations which cannot be fully explained by tectonic and other factors.

One indicator of systematic error is the value of the correlation coefficient for the b-value variation in different tectonic or geographic regions. These b-values can be estimated, using either m_b or M_S values. Kagan (1999) indicates that correlation of two sets of b-values is low; hence, conventional magnitudes may include a significant systematic error contribution.

6.3 Moment tensor errors and magnitude uncertainty

Of all the global catalogs we have considered, only the Harvard catalog estimates error for each seismic moment tensor components. In principle, these errors could be used to infer the accuracy of determining the seismic moment. However, closer analysis of these

errors (Kagan, 2000; 2002a) suggests that the random errors constitute only a part (1/3 to 1/2) of the total magnitude uncertainty. Hence, most errors should be connected with yet unknown systematic effects.

To test whether the CMT errors reflect the accuracy of magnitude determination, we investigated the dependence of the absolute value for the CMT/GS-MT magnitude differences $|\Delta m| = |m_{\rm cmt} - m_{\rm usgs}|$ on the CMT relative error (corrected Eq. 3 in Kagan 2000):

$$\epsilon = \sqrt{\sum_{i,j} E_{ij}^2 \Big/ \sum_{i,j} M_{ij}^2} \,, \tag{13}$$

where E_{ij} and M_{ij} are error and moment tensor components, respectively. We use the absolute value $|\Delta m|$ instead of Δm because the latter expression is close to zero on average, whereas the former difference represents the uncertainty of magnitude determination. For the Gaussian distribution

$$<|\Delta m|> = \frac{\sigma_{\Delta m}}{\sqrt{2\pi}} \,, \tag{14}$$

where $<>$ is a symbol of average. The regression line is

$$|\Delta m| \;=\; 0.059 + 0.37\epsilon \,, \tag{15}$$

in the ϵ-range $[0 - 0.15]$. Frohlich and Davis (1999) suggest that $\epsilon < 0.15$ corresponds to 'well-constrained' CMT solutions. Eq. 15 indicates that the CMT/GS-MT magnitude difference is strongly influenced by moment tensor component errors reported in the Harvard catalog: over the ϵ-range $[0 - 0.15]$ the average $|\Delta m|$-value increases by the factor of almost 2.0. Fig. 5 in Kagan (2002a) demonstrates that the relative error ϵ depends on earthquake size, decreasing by a factor of about 3 for large events.

We investigated a similar dependence of the $|\Delta m|$-value on the CLVD (Γ) index (or the non-DC component) for the CMT solution (see section 7.2 below). Frohlich and Davis (1999) and Kagan (2000) also suggest that the large Γ-value indicates a poor moment tensor solution. However, the dependence between the $|\Delta m|$-value and the Γ-index is weak; the former quantity increases by only about 10% over $|\Gamma|$ range $[0 - 1]$. In Fig. 6 of Kagan (2002a) the standard deviation of the Γ-index is seen as decreasing for large earthquakes. This means that for these events the absolute value of the Γ-index should decrease. A possible explanation for these features is that the Γ-index depends on the smallest eigenvalue of the moment tensor in absolute terms. This eigenvalue should not strongly influence the norm of the tensor, which is proportional to its scalar seismic moment. However, as discussed in the previous paragraph, the accuracy of the solution is generally better for large earthquakes. If almost all earthquakes are double-couples (DCs, see more below in section 7.2), the non-zero value for the smallest eigenvalue is ordinarily due to solution errors. Thus, in general the Γ-index would be relatively smaller for large earthquakes. We consider the Γ-index and the ϵ-errors in more details below.

7. Seismic moment/focal mechanism accuracy

Two catalogs we analyze in this work have either seismic moment tensor solution (CMT and GS-MT) or a first-motion DC focal mechanism solution (GS-FM). The Harvard catalog

also provides estimates of error for each tensor component. We briefly discuss the properties of these datasets as related to focal mechanism representation, to introduce the following analysis.

The seismic moment tensor is a symmetric 3×3 matrix. Its trace or the first invariant is set to be zero (Dziewonski *et al.*, 1981; Dziewonski and Woodhouse, 1983a;b; Sipkin, 1986). Hence the deviatoric tensor has five degrees of freedom. For further analysis it is convenient to separate the independent parameters into three categories: the norm of the tensor, equivalent to the scalar seismic moment, the non-DC or CLVD component of the tensor (Dziewonski and Woodhouse, 1983a, pp. 124-128; Dziewonski *et al.*, 1997), and three remaining degrees of freedom which represent orientation of a DC earthquake source – earthquake focal mechanism (Kagan, 2000; 2002a). The orientation of a DC source may be characterized by the following three quantities: a rotation angle (Φ) of counterclockwise rotation from the first DC source to the second, and a location of a rotation pole on a reference sphere – colatitude, η, and longitude, ψ (Kagan 1991c; 2000).

We discussed the scalar seismic moment or moment magnitude in the previous section. In this section, we first comment upon seismic moment tensor errors of the CMT catalog, then analyze the CLVD component for the CMT and GS-MT catalogs, and finally look at the difference of focal mechanism solutions for the two global MT catalogs above and the GS-FM catalog.

7.1 Seismic moment tensor errors

The Harvard catalog reports the standard errors for the seismic moment tensor components. We use these errors to infer the quality of the solutions and correlate them with differences of solutions in other catalogs. Since Frohlich and Davis (1999) and Kagan (2000; 2002a) discuss certain features of the errors and their relation to a tensor solution accuracy, we present here only a short outline of the error properties.

As in Dziewonski and Woodhouse (1983a;b) we use the system of coordinates (Up, South, East – r, θ, ϕ; Aki and Richards 1980, p. 118) and the 6 elements of the symmetric seismic moment tensor ordered as

$$M_{rr},\ M_{\theta\theta},\ M_{\phi\phi},\ M_{r\theta},\ M_{r\phi},\ M_{\theta\phi}\,, \tag{16}$$

or in matrix form

$$\mathbf{M} = \begin{vmatrix} M_{rr} & M_{r\theta} & M_{r\phi} \\ & M_{\theta\theta} & M_{\theta\phi} \\ & & M_{\phi\phi} \end{vmatrix}\,. \tag{17}$$

We define the relative error tensor (ε) components as (corrected equation in the first line of the section 3.2.1 in Kagan 2000)

$$\varepsilon_{ij} = < E_{ij}/M > \,. \tag{18}$$

Plots of ε_{rr} and $\varepsilon_{r\theta}$ for the 1977-2000 CMT catalog of shallow earthquakes are presented in Figs. 16a,b. As expected, the largest values of ε are for the elements $\varepsilon_{r\theta}$ and $\varepsilon_{r\phi}$; the solution for these components is unstable as the centroid depth approaches zero (Kanamori

and Given, 1981; Dziewonski *et al.*, 1981, p. 2829; Dziewonski and Woodhouse, 1983a, pp. 111-114; 1983b; Scott and Kanamori, 1985; Frohlich and Davis, 1999). Whereas the $\varepsilon_{rr}, \varepsilon_{\theta\theta}, \varepsilon_{\phi\phi}, \varepsilon_{\theta\phi}$ components scarcely depend on centroid depth, $\varepsilon_{r\theta}$ and $\varepsilon_{r\phi}$ have values 3-4 times higher for events near the Earth's surface. Only at a depth of 70 km do these $\varepsilon_{r\theta}$ and $\varepsilon_{r\phi}$ approach the other components' values (see also Eqs. 19-21 below).

For some shallow earthquakes in the Harvard catalog, no solution can be obtained for the tensor components $M_{r\theta}$ and $M_{r\phi}$. In such a case $\varepsilon_{r\theta}$ and $\varepsilon_{r\phi}$ as well as $M_{r\theta}$ and $M_{r\phi}$ are set to zero. We found 421 such $m \geq 5.3$ earthquakes in the 1977-2000 catalog (out of 10,921 total); almost all have depth set to 10 or 15 km. Two of these earthquakes have a predominantly thrust mechanism (see end of Section 3.1 for the focal mechanism determination), 334 are strike-slip events, and 85 have a normal focal mechanism. The latter two types of earthquakes are concentrated in mid-oceanic ridges and transform faults. Since these events are very shallow, this explains their poor solution convergence for $M_{r\theta}$ and $M_{r\phi}$ tensor components. Naturally, such earthquakes have been excluded from the Fig. 16b plot.

The (ε) dependence on depth (D) has been approximated by a linear regression

$$\varepsilon_{ij} \;=\; a_0 + a_1 \times D\,, \tag{19}$$

From Figs. 16a,b and similar plots, we obtain estimates for $m \geq 5.3$ earthquakes

$$a_0 = 0.029,\ 0.032,\ 0.031,\ 0.116,\ 0.122,\ 0.028\,, \tag{20}$$

and

$$a_1 = 0.00007,\ 0.00026,\ 0.00043,\ -0.00124,\ -0.00136,\ 0.00024\,, \tag{21}$$

for $\varepsilon_{rr}, \varepsilon_{\theta\theta}, \varepsilon_{\phi\phi}, \varepsilon_{r\theta}, \varepsilon_{r\phi}, \varepsilon_{\theta\phi}$, respectively.

Only minor differences (within 10-15%) have been found in error distribution for earthquakes of different focal mechanisms. However, for stronger earthquakes we find the values of a_0 to be much smaller: for $m_t = 6.0$, for instance, $a_0 \approx 0.012$ for $\varepsilon_{rr}, \varepsilon_{\theta\theta}, \varepsilon_{\phi\phi}, \varepsilon_{\theta\phi}$ components and $a_0 \approx 0.06$ for $\varepsilon_{r\theta}$ and $\varepsilon_{r\phi}$.

Relatively high values for $\varepsilon_{r\theta}$ and $\varepsilon_{r\phi}$ components mean that the total relative error ϵ in Eq. 13 would be almost completely controlled by these components for shallow events. Kanamori and Given (1981) indicate that because of large errors associated with these components for shallow earthquakes, the scalar seismic moment and moment magnitude are correlated with the fault plane dip. Therefore, errors in dip angle estimate may cause a large magnitude bias.

7.2 Non-DC (CLVD) component of seismic moment tensor

The CLVD or Γ-index has been briefly discussed in Section 6.3, where its lack of influence on the magnitude difference in two MT catalogs is explained. The non-DC component of the seismic moment tensor can be characterized by the Γ or CLVD-index:

$$\Gamma \;=\; \frac{3\sqrt{3}}{2} \times \frac{I_3}{(-I_2)^{3/2}} \;=\; \frac{3\sqrt{3}\ \lambda_1\lambda_2\lambda_3}{2\,(-\lambda_1\lambda_2 - \lambda_1\lambda_3 - \lambda_2\lambda_3)^{3/2}}\,, \tag{22}$$

where I_2 and I_3 are invariants of the deviatoric tensor (Kagan and Knopoff, 1985), and λ_i are its eigenvalues. The Γ-index ranges from -1 to 1.

We investigated the properties of the Γ-index for two catalogs, CMT and GS-MT, as well as correlation of the Γ-values for these datasets. In particular, we obtained plots of the Γ and $|\Gamma|$ dependence on earthquake time, depth, focal mechanism, and magnitude. The reason we studied the absolute Γ-value is that the Γ-distribution is largely symmetric with regard to zero. Fig. 17 displays one example of such symmetry: the dependence of the Γ-index on depth for shallow earthquakes in the CMT catalog. If the non-zero CLVD component is due to geometric or physical causes, the Γ-distribution should be asymmetric: most shallow earthquakes occur in a compressional environment, i.e., in subduction zones (Kagan, 2002a, his Table 5), thus negative Γ-index (Kuge and Lay, 1994a;b) is preferred. This Γ symmetry leads to the conclusion that for the routinely determined moment tensor, the non-DC component yields no significant information on the properties of earthquake source (see more in Kagan, 2000).

Kuge and Lay (1994b) argued that the CLVD component is largely negative for medium size ($10^{19} \geq M \geq 10^{17}$ Nm, or $6.67 \geq m \geq 5.33$) shallow earthquakes in subduction zones, where compressional stresses are predominant. They consider the CLVD determination of larger earthquakes in the CMT catalog to be biased due to use of long-period mantle waves in their solutions (see Fig. 2). To test this hypothesis, we repeated diagrams like Fig. 17 for earthquakes in this moment range, no substantial change of regression coefficients is observed in these plots.

Detailed investigations of large earthquakes, like the 1999 Chi-Chi, Taiwan (Teng *et al.*, 2001) earthquake demonstrate a significantly complex geometry of earthquake rupture. Such geometrical complexity should yield a non-zero value for the CLVD component. Since earthquakes generally exhibit a self-similar geometrical pattern, this would imply that all earthquakes should have a non-zero Γ-index.

The study of Γ-index behavior failed to find any significant dependence on time, depth, or focal mechanism. The non-DC component for the CMT catalog clearly decreases for larger earthquakes, as Fig. 6 in Kagan (2002a) demonstrates. Kuge and Lay (1994a) show the same dependence for the CMT catalog in their Fig. 1. Such an effect is observed for the GS-MT catalog, but on a much smaller scale: the $\sigma_{|\Gamma|}$ declines from 0.19 to 0.15 in the magnitude interval $[5.5 - 8.0]$. As Frohlich and Davis (1999, p. 4904) remarked, the CLVD components are significantly larger in the Harvard catalog than in the GS-MT. All the above arguments suggest that the CLVD component is not presently measured with accuracy sufficient to study its properties.

However, the strongest argument that the CLVD component in routine solutions for the seismic moment tensor results from interpretation errors and not from a physical reality is the practical lack of correlation between Γ-values in the Harvard and GS-MT catalogs. Fig. 18 displays a diagram of the Γ-indices for matched earthquakes in both catalogs. As in Fig. 8 by Frohlich and Davis (1999) the plot shows no substantial correlation between these values, a correlation which one would expect, if the CLVD component were due to properties of earthquake source. Actual correlation for shallow earthquakes is almost zero; for intermediate and deep earthquakes, similar plots show a small positive correlation between the CLVD components ($\rho \approx 0.3$ or $\rho \approx 0.4$, respectively). The magnitude threshold

increase to 6.5 or 7.0 does not modify the pattern in any significant manner.

In conclusion, the results of the Γ-index analysis can be formulated as follows. As in any physical measurement, two problems need to be solved in any quantitative investigation of the CLVD component: the size of a phenomenon to be studied and a measurement error. For the CLVD index both of these effects presently are largely unknown. The low correlation coefficient for the Γ-index in both MT catalogs suggests that its accuracy can be approximately estimated from Fig. 18: the standard error for $|\Gamma|$ is about 0.3 for the CMT dataset, whereas $\sigma_{|\Gamma|} \approx 0.15$ for the GS-MT catalog. We know even less about the distribution of the Γ-values for tectonic and other earthquakes. Kagan and Knopoff (1985) tried to estimate the index for a stochastic model of earthquake fault geometry, arguing that even if constitutive sub-events are DC sources, because of fault complexity the resulting earthquake would have a non-zero CLVD component. Similar arguments have been advanced by Kuge and Lay (1994b). Kagan and Knopoff's (1985) results imply that the CLVD index has a scale-invariant (fractal) statistical distribution and only a few percent of earthquakes may have $|\Gamma| > 0.1$. Thus, we conclude that except for a few well-studied events, almost all routine Γ determinations for individual tectonic earthquakes may turn out to be artifacts.

What can be said about studies of the CLVD components for *sets* of earthquakes, such as, for example, by Kuge and Lay (1994a;b)? In principle, if CLVD uncertainties are random and independent, for a set of n earthquakes the standard error is reduced by a factor $\sqrt{n}$, hence using many earthquakes in a particular tectonic region, we may see certain non-random properties of earthquake distribution. However, as Kuge and Lay (1994a) indicated, there may exist systematic errors for large earthquakes in the CLVD determination (see above); it is quite possible that other systematic effects are also present in the data. Thus, only after a careful analysis of uncertainties in the CLVD evaluation, could such investigations produce reliable results.

7.3 Differences between DC focal mechanisms

7.3.1 Rotation angle, Φ, between two solutions

The 3-D rotation angle, Φ, between two DC solutions represents their difference. This angle is the minimum 3-D rotation necessary to transform one DC source into another (Kagan, 1991c). We investigated the Φ-dependence on time, relative error, ϵ (13), the Γ-index (22), depth, and magnitude of matched earthquakes.

Fig. 19 displays a histogram of the rotation angle Φ distribution and its approximation by three theoretical distributions: the Cauchy, Rayleigh, and Maxwell laws. Kagan (2000) provides formulas for these distributions (see his Eqs. 7, 12, and 13, respectively) and explains why the Maxwell distribution should approximate rotation errors. As we see from the diagram, all the approximations are far from perfect: the tail of empirical distribution is heavier than that for Rayleigh and Maxwell laws; the Cauchy distribution seems to approximate the tail better, but fails to fit for small values of Φ. The latter distribution is not expected to approximate the Φ angle which is caused mostly by solution errors (see more in Kagan, 1992; 2000). The cause for the discrepancy in the Maxwell law fit is most

probably the same as for Fig. 11 – inhomogeneity of the Φ distribution. As Table 6 below demonstrates, the mean rotation angle depends on several variables. It is quite possible that there are other sources of inhomogeneity as well.

The dependence on time, illustrated in Fig. 20, testifies to a great improvement in the solution accuracy over the time-span 1980-2000. We find that the CLVD-index, the relative error, earthquake depth, and catalog time all have a major influence on the Φ-value. Table 6 illustrates these influences. We approximate the connection by a linear regression

$$< \Phi > = \; a_0 + a_1 \, (x - x_c) \, , \tag{23}$$

where x is any of the above variables, and x_c is the variable's initial value for time $T_c = 1980$.

The first six rows in Table 6 report the regression results for a CMT/GS-MT comparison. Comparing the a_0-values for different depths, we see that as in many other plots and tables discussed above, deeper earthquakes routinely have a smaller Φ-angle. Therefore, their solutions are more accurate.

For shallow matched earthquakes the dependence on the Γ-index is symmetric, so we used an absolute Γ-value for correlation. The focal mechanism accuracy seems to depend strongly on the $|\Gamma|$-value for the CMT catalog; for $|\Gamma| = 0.54$ the rotation angle increases by a factor from about 1.25 to 1.5, compared to $|\Gamma| = 0$. The $|\Gamma|$-value below of 0.54 is suggested by Frohlich and Davis (1999) as the critical level for selection of a 'well-constrained' solution (see also Kagan, 2000). However, no such regularity is observed for the $|\Gamma|$-value derived for the GS-MT catalog. Therefore, this CLVD component is a poor indicator of focal mechanism accuracy in the GS-MT case.

For the relative error ϵ (Eq. 13), we obtain an even stronger result: for $\epsilon = 0.15$ the Φ-value increases by a factor from 1.4 to 6.4, compared to $\epsilon = 0$. This ϵ-value has also been proposed by Frohlich and Davis (1999) as the critical level for selection of a 'well-constrained' solution (see also Kagan, 2000). However, a closer inspection of Φ–ϵ plots shows that relatively few earthquakes have $\epsilon \geq 0.1$.

The values of regression coefficients for catalog time T in Table 6 confirm (see Fig. 20, the regression results from the plot are entered in row 5 of Table 6) that those solutions' accuracy significantly increased during the time-span of both catalogs. Over 20 years the average Φ-value decreases by more than a factor of two.

The value of regression coefficients for the CMT/GS-FM comparison (row 7 in Table 6) reveals that the first-motion focal mechanisms have the same accuracy (or only slightly worse) as the GS-MT solutions. This result is slightly unexpected, since the first-motion focal mechanisms represent a beginning phase of earthquake rupture, whereas MT mechanism corresponds to the main phase. It is widely believed (see, for example, Scott and Kanamori, 1985; Anderson, 1988) that both focal mechanisms should differ. However, these results seem to indicate that if the difference exists, it is much smaller than the $\Phi = 15^\circ - 20^\circ$ which we observe for most angle values in Table 6.

We comment finally how magnitude influences the Φ-value (not shown in Table 6). Contrary to what may be expected, there is little change in the rotation angle as magnitude goes from 5.5 to over 8.0: for example, for shallow 1980-2000 earthquakes Φ increases only from 22.5° to 23.6°. However, if the time period is restricted to 1995-2000, the

Φ-angle decreases with magnitude from 20.0° for a $m6$ earthquake to 16.6° for a $m8$ event. This behavior change is probably caused by modified inversion procedures in the GS-MT catalog (Sipkin, 1994) around 1995: before that time focal mechanisms of large earthquakes apparently have a larger uncertainty in this catalog. Slightly stronger (and with an expected sign) influence is found for matched deep earthquakes in CMT/GS-MT catalogs: Φ goes from about 18.0° to 10.6° when magnitude changes from 5.5 to 8.0. The Φ-values decrease with magnitude increase for CMT/GS-FM comparison.

It is interesting to compare the Φ-values in Table 6 with the average Φ-value one obtains for completely random rotation of a DC source (see Fig. 4 in Kagan, 1992): $\Phi_0 = 75.2° \pm 20.9°$ is obtained by simulation. Using an expression similar to Eq. 12, we estimate that for the best determinations of focal mechanisms, uncertainty in the Φ-value for both CMT and GS-MT catalogs should be around 5° to 7°. This means that for these best focal mechanisms, the accuracy in Φ determination is only about a factor of 7 to 10 better than a random value.

7.3.2 Rotation pole distribution

Fig. 21a and 21b display the distribution of the rotation pole on a reference sphere. If the 3-D rotation is completely random, the colatitude η should be distributed according to a sinusoid in $[0° - 180°]$ range and the longitude ψ should be uniform in the range $[0° - 360°]$. In most of the diagrams, empirical histograms seem close to the theoretical distributions. In Fig. 21 we show one case where the discrepancy is large.

Table 7 summarizes rotation parameters' distribution for several depth intervals and various earthquake focal mechanisms. With regard to the Φ-value, the largest difference is displayed by normal shallow earthquakes (deep thrust events are too infrequent to draw any definite conclusion). Strike-slip shallow events have the largest difference in distribution parameters (the average and the standard deviation) from that corresponding to a random rotation. As shown in Fig. 21a the rotation pole is more often located near the horizontal plane than the random rotation distribution. Similarly, the longitude is not distributed uniformly over a sphere.

Three rotations around r, θ, and ϕ axes (Eq. 16) can be represented by the following orthogonal matrices (Altmann, 1986, p. 77)

$$\mathbf{R}_r = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos\xi & -\sin\xi \\ 0 & \sin\xi & \cos\xi \end{vmatrix}, \mathbf{R}_\theta = \begin{vmatrix} \cos\zeta & 0 & \sin\zeta \\ 0 & 1 & 0 \\ -\sin\zeta & 0 & \cos\zeta \end{vmatrix}, \mathbf{R}_\phi = \begin{vmatrix} \cos\upsilon & -\sin\upsilon & 0 \\ \sin\upsilon & \cos\upsilon & 0 \\ 0 & 0 & 1 \end{vmatrix}, \tag{24}$$

where ξ, ζ, and υ are rotation angles. Comparing (24) with (17) we see that $\epsilon_{r\theta}$ and $\epsilon_{r\phi}$ errors cause rotation of the tensor around East-West (E-W) or South-North (S-N) axes, respectively.

As we see from Figs. 16a,b, for shallow earthquakes $\epsilon_{r\theta}$ and $\epsilon_{r\phi}$ exceed all other moment component errors. Three principal focal mechanisms (thrust, strike-slip, and normal) can be represented by the following diagonal matrices

$$\text{diag}[1 \;\; 0 \;\; -1], \quad \text{diag}[0 \;\; 1 \;\; -1], \quad \text{and} \quad \text{diag}[-1 \;\; 0 \;\; 1]\,. \tag{25}$$

Rotation by the $\mathbf{R}_r$ matrix would change only a strike of a focal mechanism without changing its type, thus the first term in (25) matrices would remain constant, whereas $M_{\theta\theta}$, $M_{\theta\phi}$, and $M_{\phi\phi}$ undergo appropriate transformations.

From (24) and (25) expressions we hypothesize that for shallow earthquakes the error rotation pole would concentrate in a horizonal plane. Simulations confirm this conjecture. As in Kagan (2000), we simulate rotation of an individual focal mechanism (25), using either rotations in (24) or rotations induced by random errors in tensor components.

As another technique we use moment component errors in the CMT catalog to simulate average mechanism rotation for all earthquakes in 0-35 km range, as well as rotation for three classes of focal mechanisms. In Fig. 22 we demonstrate the distribution of the rotation pole if the rotation is caused by Gaussian distributed errors with the standard error listed in the CMT catalog. The diagrams show that for shallow (0-35 km) events, the rotation pole is more likely to be near the horizontal plane, corresponding to colatitude $\eta \approx 90°$, especially so for strike-slip earthquakes. The rotations in Fig. 22 are not fully analogous to those in Fig. 21: first, the tensor component errors in the CMT catalog do not represent the whole uncertainty (Kagan, 2000); second, solution errors in both CMT/GS-MT catalogs cause rotations in Fig. 21.

Both eigenvalues for strike-slip earthquakes are influenced by S-N and E-W rotations. Thus, these events should be more strongly rotated due to $\epsilon_{r\theta}$ and $\epsilon_{r\phi}$ errors. Since these rotations should not significantly change eigenvalues, the CLVD index would not be affected by them. Table 1 in Kagan (2000) confirms these conjectures: the rotation angle Φ for a strike-slip earthquake is larger than that for other focal mechanisms, whereas the Γ-index variance is the smallest one.

Distribution of the longitude is more complex: our simulations show that it largely depends on the strike distribution of focal mechanisms for considered earthquakes. A more complete analysis of rotation pole pattern would require more detailed investigations. Analytical representation of rotations induced by tensor component errors would facilitate such an analysis. Xu and Grafarend (1996) and Xu (1999) investigated the geometry of the eigenspectra for general and constrained second-rank tensors, although the solution for the DC tensor in the representation M, Φ, η, and ψ (see the beginning of this section) is still, as far as we know, unavailable.

8. Discussion

Our aim has been to analyze the whole range of catalog uncertainties: catalog completeness, earthquake time, location, magnitude errors, as well as seismic moment tensor solution accuracy. The Harvard CMT catalog clearly is the most complete and detailed in existence; thus, we attempted to analyze its accuracy as much as was feasible.

The major tool of our investigations has been to compare parameters for earthquakes matched in several catalogs. Although some catalogs provide estimates of parameter uncertainties, in origin time and location errors, these quantities show only an internal discrepancy of solutions. Each such solution has systematic effects which may exceed the reported random error values. Moreover, most catalogs lack reported uncertainties for earthquake magnitude or for most important parameters of moment tensor solutions.

What is the source of earthquake parameter differences we found? Some disagreements

can be traced to a different frequency range of seismograms used in data interpretation. Magnitudes and scalar seismic moments depend on the frequency, measure different earthquake properties, and thus cannot be expected to be identical. In principle seismic moment should be estimated at zero frequency (Aki and Richards, 1980); however, in practical evaluations (Dziewonski *et al.*, 1981; Dziewonski and Woodhouse, 1983a;b; Sipkin, 1986; Sipkin *et al.*, 2002) low frequency waves are used. Since the period of these waves is occasionally comparable or even less than the rupture time of the greatest earthquakes, an estimation bias is unavoidable.

The other source of parameter differences are various interpretation techniques and assumptions used in seismogram processing. Since these particulars of technique are not fully recorded, their accounting is difficult. And, finally, difference is caused by random errors which are the effect of model deficiencies and insufficient knowledge of the Earth's structure.

An important issue in comparing catalogs is degree of statistical independence in data entries. Clearly, earthquake parameters listed in datasets are not fully independent: for instance, the CMT and GS-MT catalogs usually initiate solution construction after receiving a message from the PDE or ISC catalog compilers. Thus, one should expect that moment tensor catalogs would list almost all moderate or large earthquakes, available in the PDE or ISC catalogs.

Catalog compilers also use sets of seismograms which at least partially coincide; the Earth structure parameters are also not independent and may share similar systematic effects. Therefore, differences in parameters of matched earthquakes may under-estimate real earthquake errors. However, despite these qualifications, our statistical analysis provides at least a lower bound for possible uncertainty values. Future investigations may improve the accuracy of such estimates.

Another important issue in catalog accuracy analysis is the presence of regional or local variations in earthquake parameters or seismicity patterns. These variations may result from different techniques used in seismogram processing, or have an underlying physical cause. For example, if it can be shown that earthquake distribution parameters obtained while using different techniques or different datasets have no significant correlation over regional or local scales, this finding would support the spurious character of these features (see the end of Section 6.2).

9. Conclusions

We analyze four worldwide earthquake catalogs to infer their completeness and accuracy of earthquake parameters. We confined our studies mostly to the Harvard CMT catalog, especially its moment tensor solutions. The results of investigations can be summarized as follows:

1. We have devised more rigorous procedures for testing the completeness of catalogs and accurately estimating lower magnitude thresholds.

2. The CMT catalog is reasonably complete, with the magnitude threshold changing from about 5.7 in the first years to the present value of about 5.4. The threshold also varies with earthquake depth, tectonic provinces, and earthquake focal mechanism.

3. Seismic moment tensor catalogs based on low frequency seismic data (like the CMT)

miss some earthquakes, even quite large ones, when they occur close in time to others. These could seriously affect the results of earthquake clustering and aftershock studies. Additional work is needed to quantify this feature of catalogs.

4. Origin time and location differences between MT and conventional earthquake catalogs can be explained by different Earth's structure adopted in interpretation, as well as by random errors and extended source properties of large earthquakes. The latter feature is well approximated by regular earthquake scaling relations.

5. Accuracy of magnitude evaluation estimated for moment magnitude uncertainty is on the order 0.05–0.09, depending on earthquake depth, catalog time, and magnitude. For the CMT catalog, the magnitude errors strongly correlate with seismic tensor component errors, allowing an additional method to estimate this uncertainty.

6. Conversion of conventional magnitudes into moment magnitude leads to magnitude errors which are by a factor of three to four higher than errors in the MT catalogs. The saturation effect makes these magnitudes even less reliable for estimating the size of large earthquakes.

7. No significant correlation between non-DC components of CMT and GS-MT solutions suggests that routinely determined CLVD-values do not reliably indicate deviation of earthquake focal mechanisms from a standard fault model.

8. We analyzed how seismic tensor component errors reported in the CMT catalog influence major parameters of earthquake focal mechanisms. Due to large errors for near Earth surface earthquakes, their solutions have a larger uncertainty.

9. Focal mechanism error as measured by a 3-D rotation angle depends on catalog time, earthquake depth, magnitude, CLVD-index, tensor component errors. The angle varies from $15^\circ - 20^\circ$ to $5^\circ - 7^\circ$ for the best solutions.

10. The distribution of a 3-D rotation pole for shallow earthquakes exhibits features explainable by large tensor component errors. More thorough analysis is needed to explore regularities of 3-D rotation of earthquake focal mechanisms.

Acknowledgments

I appreciate partial support from the National Science Foundation through grant EAR 00-01128 and from the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-8920136 and USGS Cooperative Agreements 14-08-0001-A0899 and 1434-HQ-97AG01718. I thank D. D. Jackson and F. Schoenberg for stimulating discussions and help in formulating some of my results. The comments by the reviewers Göran Ekström and Cliff Frohlich as well as by the Editor Ken Creager have significantly improved the presentation. Publication 691, SCEC.

References

Aki, K., 1965. Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits, *Bull. Earthquake Res. Inst. Tokyo Univ.*, **43**, 237-239.

Aki, K. and Richards, P., 1980. *Quantitative Seismology*, W. H. Freeman, San Francisco, 2 Vols, 557 and 373 pp.

Altmann, S. L., 1986. *Rotations, Quaternions and Double Groups*, Clarendon Press, Oxford, pp. 317.

Anderson, H., 1988. Comparison of centroid-moment tensor and 1st motion solutions for western Mediterranean earthquakes, *Phys. Earth Planet. Inter.*, **52**, 1-7.

Bird, P., Kagan, Y. Y. and Jackson, D. D., 2002. Plate tectonics and earthquake potential of spreading ridges and oceanic transform faults, in: *Plate Boundary Zones, AGU Monograph*, eds. S. Stein and J. T. Freymueller, 203-218 (doi: 10/1029/030GD12).

Bol'shev, L. N., 1963. Asymptotic Pearson transformations, *Theory Prob. Appl.*, **8**, 121-146 (English translation).

Christophersen, A., 2000. *The Probability of a Damaging Earthquake Following a Damaging Earthquake*, PhD thesis, Victoria Univ. of Wellington, New Zealand, pp. 111.

Deemer, W. L. and Votaw, D. F., 1955. Estimation of parameters of truncated or censored exponential distributions, *Ann. Math. Stat.*, **26**, 498-504.

Dziewonski, A. M., and D. Anderson, 1981. Preliminary reference Earth model (PREM), *Phys. Earth Planet. Inter.*, **25**, 297-356.

Dziewonski, A. M., Chou, T.-A. and Woodhouse, J. H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, **86**, 2825-2852.

Dziewonski, A. M. and Woodhouse, J. H., 1983a. Studies of the seismic source using normal-mode theory, in: *Earthquakes: Observation, Theory and Interpretation*, Proc. Int. School Phys. "Enrico Fermi", Course LXXXV, eds H. Kanamori and E. Boschi, North-Holland Publ., Amsterdam, 45-137.

Dziewonski, A. M. and Woodhouse, J. H., 1983b. An experiment in systematic study of global seismicity: centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981, *J. geophys. Res.*, **88**, 3247-3271.

Dziewonski, A. M., Ekström, G. and Nettles, M., 1997. Harvard centroid-moment tensor solutions 1976-96: Significance of the non-double-couple component, in: *Proceedings of the 4th International Symposium on Rockbursts and Seismicity in Mines*, Krakow, Poland, 11-14 August 1997, edited by S.J. Gibowicz, S. Lasocki, Rotterdam, Brookfield, VT, A.A. Balkema, pp. 3-16.

Dziewonski, A. M., Ekström, G. and Maternovskaya, N. N., 1999. Centroid-moment tensor solutions for October-December, 1998. *Phys. Earth Planet. Inter.*, **115**, 1-16.

Dziewonski, A. M., Ekström, G. and Maternovskaya, N. N., 2001. Centroid-moment tensor solutions for July-September 2000, *Phys. Earth Planet. Inter.*, **124**, 9-23. The catalog is available by FTP anonymous to saf.harvard.edu, cd /CMT.

Ekström, G. and Dziewonski, A. M., 1988. Evidence of bias in estimation of earthquake size, *Nature*, **332**, 319-323.

Ekström, G. and Nettles, M., 1997. Calibration of the HGLP seismograph network and centroid-moment tensor analysis of significant earthquakes of 1976, *Phys. Earth Planet. Inter.*, **101**, 219-243.

Engdahl, E. R., van der Hilst, R. D. and Buland, R. P., 1998. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Amer.*, **88**, 722-743.

Frohlich, C., 2001. Display and quantitative assessment of distributions of earthquake focal mechanisms, *Geophys. J. Int.*, **144**, 300-308.

Frohlich, C. and Davis, S. D., 1999. How well constrained are well-constrained T, B, and P axes in moment tensor catalogs?, *J. Geophys. Res.*, **104**, 4901-4910.

Fukuyama, E. and Dreger, D. S., 2000. Performance test of an automated moment tensor determination system for the future "Tokai" earthquake, *Earth, Planets and Space (EPS)*, **52**, 383-392.

Hanks, T. C. and Kanamori, H., 1979. A moment magnitude scale, *J. Geophys. Res.*, **84**, 2348-2350.

Harte, D. and Vere-Jones, D., 1999. Differences in coverage between the PDE and New Zealand local earthquake catalogues, *New Zealand J. Geol. Geoph.*, **42**, 237-253.

Helffrich, G. R., 1997. How good are routinely determined focal mechanisms? Empirical statistics based on a comparison of Harvard, USGS and ERI moment tensors, *Geophys. J. Int.*, **131**, 741-750.

Hileman, J. A., Allen, C. R. and Nordquist, J. M., 1973. *Seismicity of the Southern California Region, 1 January 1932 to 31 December 1972*, Cal. Inst. Technology, Pasadena.

International Seismological Centre (ISC), *Bull. Intern. Seismol. Centre*, **32**, 1995.

Kagan, Y. Y., 1991a. Seismic moment distribution, *Geophys. J. Int.*, **106**, 123-134.

Kagan, Y. Y., 1991b. Likelihood analysis of earthquake catalogues, *Geophys. J. Int.*, **106**, 135-148.

Kagan, Y. Y., 1991c. 3-D rotation of double-couple earthquake sources, *Geophys. J. Int.*, **106**, 709-716.

Kagan, Y. Y., 1992. Correlations of earthquake focal mechanisms, *Geophys. J. Int.*, **110**, 305-320.

Kagan, Y. Y., 1995. Magnitude-frequency distribution in the European-Mediterranean earthquake regions – Comment, *Tectonophysics*, **245**, 101-105.

Kagan, Y. Y., 1997. Seismic moment-frequency relation for shallow earthquakes: Regional comparison, *J. Geophys. Res.*, **102**, 2835-2852.

Kagan, Y. Y., 1999. Universality of the seismic moment-frequency relation, *Pure Appl. Geoph.*, **155**, 537-573.

Kagan, Y. Y., 2000. Temporal correlations of earthquake focal mechanisms, *Geophys.*

J. Int., **143**, 881-897.

Kagan, Y. Y., 2002a. Seismic moment distribution revisited: I. Statistical results, *Geophys. J. Int.*, **148**, 520-541.

Kagan, Y. Y., 2002b. Modern California earthquake catalogs and their comparison, *Seism. Res. Lett.*, **73**(6), 921-929.
http://scec.ess.ucla.edu/~ykagan/calcat_index.html.

Kagan, Y. Y. and Knopoff, L., 1985. The first-order statistical moment of the seismic moment tensor, *Geophys. J. Roy. astr. Soc.*, **81**, 429-444.

Kanamori, H., 1977. The energy release in great earthquakes, *J. Geophys. Res.*, **82**, 2981-2987.

Kanamori, H. and Given, J. W., 1981. Use of long-period surface waves for rapid determination of earthquake-source parameters, *Phys. Earth Planet. Inter.*, **27**, 8-31.

Kubo, A., E. Fukuyama, H. Kawai, and K. Nonomura, 2002. NIED seismic moment tensor catalogue for regional earthquakes around Japan: quality test and application, *Tectonophysics*, **356**, 23-48.

Kuge, K., 1992. Systematic difference in the ISC body-wave magnitude–seismic moment relationship between intermediate and deep earthquakes, *Bull. Seismol. Soc. Amer.*, **82**, 819-835.

Kuge, K. and Lay, T., 1994a. Data-dependent non-double-couple components of shallow earthquake source mechanisms – effects of waveform inversion instability, *Geophys. Res. Lett.*, **21**, 9-12.

Kuge, K. and Lay, T., 1994b. Systematic non-double-couple components of earthquake mechanisms – the role of fault zone irregularity, *J. Geophys. Res.*, **99**, 15,457-15,467.

Molchan, G., Kronrod, T. and Panza, G. F., 1997. Multi-scale seismicity model for seismic risk, *Bull. Seismol. Soc. Amer.*, **87**, 1220-1229.

Preliminary Determination of Epicenters, (PDE), Monthly Listings, 1999. U.S. Dept. Interior/Geol. Survey, Nat. Earthquake Inform. Center, January, pp. 47, WEB: http://wwwneic.cr.usgs.gov/neis/data_services/ftp_files.html.

Needham, R. E., 1988. *Catalog of First Motion Focal Mechanisms, 1986-1987*, **Vols 1,2,3**, *Open-file report 88-0556-A,B,C*, pp. 236, 231, 298, U. S. Geological Survey.

Ogata, Y., 1998. Space-time point-process models for earthquake occurrences, *Ann. Inst. Statist. Math.*, **50**(2), 379-402.

Ogata, Y., 1999. Seismicity analysis through point-process modeling: A review, *Pure Appl. Geophys.*, **155**, 471-507.

Pasyanos, M. E., Dreger, D. S. and Romanowicz, B., 1996. Toward real-time estimation of regional moment tensors, *Bull. Seismol. Soc. Amer.*, **86**, 1255-1269.

Patton, H. J., 2001. Regional magnitude scaling, transportability, and M_s:m_b discrimination at small magnitudes, *Pure Appl. Geophys.*, **158**, 1951-2015.

Pondrelli, S., Morelli, A., Ekström, G., Mazza, S., Boschi, E., Dziewonski, A. M., 2002. European-Mediterranean regional centroid-moment tensors: 1997-2000, *Phys. Earth*

Planet. Inter., **130**, 71-101.

Röhm, A. H. E., Trampert, J., Paulssen, H. and Snieder, R. K., 1999. Bias in reported seismic arrival times deduced from the ISC Bulletin, *Geophys. J. Int.*, **137**, 163-174.

Scott, D. R. and Kanamori, H., 1985. On the consistency of moment tensor source mechanisms with first-motion data, *Phys. Earth Planet. Inter.*, **37**, 97-107.

Sipkin, S. A., 1986. Estimation of earthquake source parameters by the inversion of waveform data: global seismicity, 1981-1983, *Bull. Seismol. Soc. Amer.*, **76**, 1515-1541.

Sipkin, S. A., 1994. Rapid determination of global moment-tensor solutions, *Geophys. Res. Lett.*, **21**, 1667-1670.

Sipkin, S. A., Bufe, C. G. and Zirbes, M. D., 2002. Moment-tensor solutions estimated using optimal filter theory: global seismicity, 2000, *Phys. Earth Planet. Inter.*, **130**, 129-142.

Smith, G. P. and Ekström, G., 1995. Using travel-time and waveform data to study the earthquake source, *Eos Trans. AGU*, **76**(46), Fall AGU Meet. Suppl., (abstract), p. F389.

Smith, G. P. and Ekström, G., 1997. Interpretation of earthquake epicenter and CMT centroid locations, in terms of rupture length and direction, *Phys. Earth Planet. Inter.*, **102**, 123-132.

Stephens, M. A., 1974. EDF statistics for goodness of fit and some comparisons, *J. Amer. Statist. Assoc. (JASA)*, **69**, 730-737.

Storchak, D. A., Bird, A. L. and Adams, R. D., 2000. Discrepancies in earthquake location between ISC and other agencies, *J. Seismology*, **4**, 321-331.

Teng, T. L., Tsai, Y. B. and Lee, W. H. K., 2001. Preface to the 1999 Chi-Chi, Taiwan, earthquake dedicated issue, *Bull. Seismol. Soc. Amer.*, **91**, 893-894; Collection of papers, 893-1394.

Vere-Jones, D., Robinson, R. and Yang, W. Z., 2001. Remarks on the accelerated moment release model: problems of model formulation, simulation and estimation, *Geophys. J. Int.*, **144**, 517-531.

Wiemer, S. and Wyss, M., 2000. Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western United States, and Japan, *Bull. Seismol. Soc. Amer.*, **90**, 859-869.

Xu, P. L., 1999. Spectral theory of constrained second-rank symmetric random tensors, *Geophys. J. Int.*, **138**, 1-24.

Xu, P. L. and Grafarend, E., 1996. Statistics and geometry of the eigenspectra of three-dimensional second-rank symmetric random tensors, *Geophys. J. Int.*, **127**, 744-756.

Zhu, L. P. and Helmberger, D. V., 1996. Advancement in source estimation techniques using broadband regional seismograms, *Bull. Seismol. Soc. Amer.*, **86**, 1634-1641.

List of captions

Fig. 1. Time distribution of earthquake numbers in the 1980/1/1–2001/7/31 GS-MT catalog: $m \geq 5.6$, $m \geq 5.8$, $m \geq 6.0$, $m \geq 6.2$. Average number of events is shown by straight horizontal lines. The numbers of weaker ($m \leq 6.0$) shallow earthquakes stabilize only after 1995.

Fig. 2. Magnitude-frequency relation for the Harvard 1977/1/1–2000/12/31 catalog. Solid lines for all earthquakes, where the solutions are based on body waves; dashed lines for earthquakes in which long-period mantle waves have also been used in inversions. G-R approximations are shown by dotted lines. To avoid overlapping of the curves we multiply the number of intermediate events by 0.1 and the number of deep earthquakes by 0.01. Each of these techniques (body *vs* mantle waves) may have particular difficulties in their solutions (*cf.* Kuge and Lay, 1994a), thus the transition from one algorithm to another may introduce different biases in determination of the scalar seismic moment, M.

Fig. 3. Dependence of the β-value estimate on the moment magnitude cutoff, m_v, for the Harvard 1977/1/1–2000/12/31 catalog. Solid line – shallow earthquakes (0-70 km); dashed line – intermediate earthquakes (70-300 km); dotted line – deep earthquakes (300-700 km). Curves stabilize around $m_v \geq 5.5$ which roughly correspond to the threshold value. The β-values at the plateau – $\beta \approx \frac{2}{3}$ correspond to the traditional G-R b-value equal to 1.0.

Fig. 4. Dependence of the significance level, α, for rejection of the Pareto (G-R) distribution on the moment magnitude cutoff, m_v, for the Harvard 1977/1/1–2000/12/31 catalog. Solid line – shallow earthquakes (0-70 km); dashed line – intermediate earthquakes (70-300 km); dotted line – deep earthquakes (300-700 km). The curves increase sharply at some magnitude cutoff value. Selecting an appropriate significance level α, we determine the threshold m_t-value.

Fig. 5. Temporal distribution of earthquake magnitudes in the 1980-2001/7/31 GS-MT catalog for which there is no match in the Harvard catalog. Depth range is 0-700 km. Most of earthquakes missing from the Harvard catalog are associated with complex events which were interpreted differently in the GS-MT list.

Fig. 6. Distribution of time intervals between earthquakes in the Harvard catalog, 1977-2000. The maximum separation between centroids is 250 km. Solid line – $6.0 > m \geq 5.5$; dashed line – $6.5 > m \geq 6.0$; dash-dotted line – $7.0 > m \geq 6.5$; dotted line – $7.5 > m \geq 7.0$; circles – $8.0 > m \geq 7.5$; '$\times$' and solid line – $8.5 > m \geq 8.0$; where m above is the magnitude of the first earthquake in a sequence. The shift of curves to longer time intervals for larger earthquakes may be explained by their longer coda.

Fig. 7. Plot of time difference (Δt) between the origin time in the PDE catalog and the

Harvard CMT centroid time for shallow (0-70 km) earthquakes. Dashed line – linear approximation; solid line – quadratic approximation. Results of both regressions are written at the top of the plot: ρ is the coefficient of correlation, σ – standard error, ε_{max} – maximum difference, n – number of pairs.

Fig. 8. Plot of binned time differences between the origin time in the PDE catalog and the Harvard CMT centroid time for shallow (0-70 km) earthquakes. We exclude the last bin that contains only three earthquakes (see Fig. 7). Solid line – approximation from Smith and Ekström (1995). Circles – average difference; pluses and stars – average difference and $\pm$ standard error, respectively.

Fig. 9. Plot of dependence of average epicentral distance between CMT centroid and PDE epicenter for shallow (0-70 km) earthquakes on CMT moment magnitude. Circles – average difference; pluses and stars – average difference and $\pm$ standard error, respectively. Solid line – the power-law approximation from Smith and Ekström (1997), dashed line – approximation by a quadratic polynomial.

Fig. 10. Plot of dependence of average epicentral distance between PDE and ISC epicenters for shallow (0-70 km) earthquakes on m_b PDE magnitude. Time limits are 1971/1/1–1998/12/31. Circles – average difference; pluses and stars – average difference and $\pm$ standard error, respectively. The last three points in the plot (for $m_b \geq 6\frac{7}{8}$) are based on a few (17) earthquake pairs, hence they are less reliable.

Fig. 11. Histogram of epicentral distances for pairs of matched shallow (0-70 km) earthquakes ($m_b \geq 5.0$) in the PDE and ISC catalogs. Time limits are 1971/1/1–1998/12/31. The solid line is an approximation by the Rayleigh distribution having the same mean as the empirical distribution. Poor fit between the theoretical and experimental distributions may be due to a significant inhomogeneity of location errors.

Fig. 12. Dependence of the moment magnitude difference Δm between two catalogs (Harvard and GS-MT) on the Harvard moment magnitude ($m \geq 6.0$). Dashed line – linear approximation (Eq. 8); solid line – quadratic approximation (Eq. 9). Results of both regressions are written at the top of the plot: ρ is coefficient of correlation, σ – standard error, ε_{max} – maximum difference, n – number of pairs.

Fig. 13. Dependence of the magnitude difference Δm between two catalogs (Harvard and PDE-M_S) on the Harvard moment magnitude. Dashed line – linear approximation; solid line – quadratic approximation. Results of both regressions are written at the top of the plot, see Fig. 12 for format.

Fig. 14. Body-wave magnitude m_b from the PDE catalog as a function of the Harvard moment magnitude. Shallow earthquakes (0-70 km) 1977-2000 are used in comparison. Dashed line – linear approximation; solid line – quadratic approximation. Results of both regressions are written at the top of the plot, see Fig. 12 for format.

In regression analysis results, shown in the plot, m_r is taken 5.3 (see Eqs. 8 and 9).

Fig. 15. Dependence of the magnitude difference for CMT/GS-MT catalogs and its standard deviation on catalog time. All matched $m_{\rm cmt} \geq 6$ earthquakes (depth range 0-700 km) are included. Circles – average difference; pluses and stars – average difference and $\pm$ standard error, respectively. Improvement in magnitude determination accuracy is clearly seen in the plot.

Figs. 16a,b. Dependence of the Harvard seismic moment tensor errors on depth.
a) Relative error ε_{rr};
b) Relative error $\varepsilon_{r\theta}$.
Dashed line – linear approximation, a_0 corresponds to depth $D = 0$ km; solid line – quadratic approximation. Results of both regressions are written at the top of the plot, see Fig. 12 for format. Whereas the relative error ε_{rr} does not significantly depend on depth, $\varepsilon_{r\theta}$ strongly decays with depth.

Fig. 17. Dependence of the CLVD index Γ for shallow $m \geq 5.3$ earthquakes in the 1977-2000 CMT catalog on depth.
Dashed line – linear approximation, a_0 corresponds to depth $D = 0$ km; solid line – quadratic approximation. The results of both regressions are written at the top of the plot, see Fig. 12 for format.

Fig. 18. Correlation between the CLVD indices (Γ) for matched shallow $m \geq 5.3$ earthquakes in the CMT/GS-MT catalogs.
Dashed line – linear approximation; solid line – quadratic approximation. The results of both regressions are written at the top of the plot, see Fig. 12 for format.

Fig. 19. Histogram of the 3-D rotation angles (Φ) between matched pairs of CMT/GS-MT focal mechanism DC-solutions for shallow $m \geq 6$ earthquakes. Solid line is approximation by the Rayleigh distribution. Dash-dotted line is the Maxwell distribution approximation. Dashed line is the Cauchy distribution approximation. All the theoretical distributions have the same mean $<\Phi>$ as the empirical distribution. Also shown in the plot are the rotation angle standard deviation σ_Φ, the Φ median, and the number of pairs N. Poor fit between the theoretical and experimental distributions may be due to a significant inhomogeneity of focal mechanism errors.

Fig. 20. Dependence of the 3-D rotation angles (Φ) between matched pairs of CMT/GS-MT focal mechanism DC-solutions for $m \geq 6$ earthquakes on catalog time. Dashed line – linear approximation, a_0 corresponds to year $T_c = 1980$. Results of both regressions are written at the top of the plot, see Fig. 12 for format.

Fig. 21a,b. Histogram of the 3-D rotation pole for rotation between matched pairs of CMT/GS-MT focal mechanism DC-solutions for shallow (0-35 km) $m \geq 5.3$ strike-slip earthquakes. Statistical parameters of the distributions are written at the top of the plot, see Fig. 19 for format.

a) Colatitude (η). Dashed line is the sinusoidal distribution approximation. Zero of colatitude corresponds to the vector of the rotation pole pointing down (Kagan 1991c, Eq. 18).
b) Longitude (ψ). Dashed line is the uniform distribution approximation.

Fig. 22a,b,c. Distribution of rotation poles over a reference sphere. Earthquake depth range 0-35 km. CMT tensor component errors are used to simulate rotation of DC focal mechanisms.
a) Thrust earthquakes;
b) Strike-slip earthquakes;
c) Normal earthquakes.
Rotation poles are concentrated near the horizontal plane ($\eta = 90^\circ$), especially so for strike-slip earthquakes for which both eigenvalues are influenced by rotations caused by errors.

TABLE 1. Magnitude threshold m_t for the Harvard CMT catalog

		Shallow 0-70 km		Intermediate 70-300 km		Deep 300-700 km	
#	Years	α_u	α	α_u	α	α_u	α
1	1977-2001	5.71	5.72	5.28	5.28	5.33	5.34
2	1977-1982	5.70	5.68	5.39	5.37	5.31	5.34
3	1982-1987	5.47	5.77	5.25	5.25	5.22	5.22
4	1987-1992	5.40	5.40	5.25	5.26	5.24	5.27
5	1992-1997	5.44	5.40	5.58	5.65	5.42	5.42
6	1997-2001	5.27	5.34	5.15	5.15	5.30	5.30

Magnitude threshold calculation is based on the Kolmogorov test for the significance level (α) equal 0.1. The significance level α_u is calculated assuming the universal value for β: $\beta = 0.667$, $\beta = 0.64$, and $\beta = 0.62$, for shallow, intermediate, and deep earthquakes, respectively (Kagan, 2002a). The significance level α calculated by estimating $\hat{\beta}$ (Kagan, 2002a, Eq. 22) and fitting the empirical distribution with the Pareto (G-R) law.

TABLE 2. Magnitude threshold m_t for the Harvard CMT catalog in Flinn-Engdahl seismic regions

#	FE regions	Earthquakes			
		All	Thrust	Strike-slip	Normal
1	Subduction	5.72	5.67	6.06	5.30
2	Collision	5.32	5.17	5.32	5.16
3	Continental	5.21	4.85	5.08	5.20
4	Oceanic	5.64	5.30	5.73	5.40
5	Other	5.39	5.37	5.36	5.36

Magnitude threshold calculation is based on the Kolmogorov test for the significance level (α) equal 0.1. The level α is calculated by estimating $\hat{\beta}$ (Kagan, 2002a, Eq. 22) and fitting the empirical distribution with the Pareto (G-R) law.

TABLE 3. Inter-earthquake time distribution parameter $\tau_{0.1}$

#	Catalogs	Magnitude	$\tau_{0.1}$ days
1	ISC	m_b	0.014
2	PDE	m_b	0.017
3	PDE	M_S	0.100
4	CMT	m_w	0.133
5	GS-MT	m_w	0.450

$\tau_{0.1}$ is the inter-earthquake time interval for lower 10% distribution for the first earthquake in $6.0 > m \geq 5.5$ magnitude range.

TABLE 4. Origin time differences (Δt) in catalogs

#	Cat.	Year	Depth km	m/m	n	a_0	a_1	$\sigma_{\Delta t}$
							sec	
1	c/p	77-00	0-70	m_w/M_S	2351	4.00	7.80	3.74
2	c/p	77-00	0-70	m_w/m_b	2463	3.90	7.98	3.83
3	c/p	77-00	70-300	m_w/m_b	461	3.64	6.26	3.08
4	c/p	77-00	300-700	m_w/m_b	228	5.16	3.84	2.67
5	c/p	77-00	0-700	m_w/m_b	3235	3.90	7.43	3.73
6	c/i	77-98	0-700	m_w/m_b	2963	3.66	7.43	3.90
7	p/i	71-98	0-70	m_b/m_b	27807	−0.16	0.13	1.82
8	p/i	71-98	70-300	m_b/m_b	7054	−0.27	0.17	1.66
9	p/i	71-98	300-700	m_b/m_b	2451	0.14	0.00	1.54

Dependence of origin time differences on magnitude (the first magnitude shown in column 5). Catalogs' abbreviations are: 'c' – for the Harvard CMT catalog, 'p' – for the PDE catalog, and 'i' – for the ISC catalog. 'm/m' – magnitudes, used in comparison. 'n' – the number of matched pairs in each of two catalogs. We match $m \geq 6$ CMT earthquakes with PDE and ISC events. When comparing PDE/ISC earthquakes, we use $m_b \geq 5$ events. Coefficients a_0, a_1, and standard error $\sigma_{\Delta t}$ (columns 7-9) are calculated for linear regression (similar to Eq. 8). Coefficient a_0 is calculated for rows 1-6 using $m_r = 6$ and for rows 7-9 using $m_r = 5$.

TABLE 5. Magnitude differences in catalogs

#	Cat.	Year	Depth km	m/m	n	a_0	a_1	a_2	$\sigma_{\Delta m}$
						$\times 1000$			
1	c/g	80-00	0-35	m_w/m_w	992	1	5	52	124
2	c/g	80-00	0-70	m_w/m_w	1389	2	−17	64	116
3	c/g	80-00	70-300	m_w/m_w	304	−19	−28	52	78
4	c/g	80-00	300-700	m_w/m_w	133	−18	19	23	68
5	c/g	95-00	0-70	m_w/m_w	707	31	18	22	83
6	c/p	77-00	0-35	m_w/m_b	1756	437	232	166	264
7	c/p	77-00	0-70	m_w/m_b	2463	403	280	141	264
8	c/p	77-00	70-300	m_w/m_b	461	314	265	136	226
9	c/p	77-00	300-700	m_w/m_b	228	437	302	48	207
10	c/p	77-00	0-70	m_w/M_S	2351	312	−375	157	201
11	g/p	95-00	0-700	m_w/m_b	591	393	249	132	256
12	c/p	95-00	0-700	m_w/m_b	765	457	161	160	265
13	g/p	95-00	0-700	m_w/M_S	470	292	−286	95	264
14	c/p	95-00	0-700	m_w/M_S	601	360	−340	130	236
15	c/i	77-98	0-70	m_w/m_b	2220	460	311	138	258
16	c/i	77-98	0-700	m_w/m_b	2959	446	316	130	248
17	g/i	80-98	0-70	m_w/m_b	1225	353	437	97	245
18	g/i	80-98	0-700	m_w/m_b	1675	349	447	77	234
19	p/i	71-98	0-70	m_b/m_b	27862	91	62	44	122
20	p/i	71-98	0-70	M_S/m_b	10315	373	508	32	275
21	p/p	71-00	0-70	M_S/m_b	9836	306	477	30	274

Dependence of magnitude differences on magnitude (the first magnitude shown in column 5). Catalogs' abbreviations are: 'c' – for the Harvard CMT catalog, 'g' – for the GS-MT catalog, 'p' – for the PDE catalog, and 'i' – for the ISC catalog. 'm/m' – magnitudes, used in comparison. 'n' – the number of matched pairs in each of two catalogs. Parameters a_0, a_1, a_2, and σ are multiplied by 1000 to save display space. We match $m \geq 6$ CMT earthquakes with GS-MT, PDE and ISC events. When comparing PDE/ISC earthquakes, we use $m_b \geq 5$ or $M_S \geq 5$ events. Coefficient a_0 (column 7) is calculated for all entries using $m_r = 6$ (see Eq. 9).

TABLE 6. Dependence of 3-D rotation angle Φ on CLVD index, relative error, and catalog time

#	Cat. Type	Depth km	n	$\lvert\Gamma\rvert$			ϵ		T	
				a_0	a_1	σ_Φ	a_0	a_1	a_0	a_1
1	c/g	0-35	992	22.4	11.7	16.1	17.0	163	37.3	−0.97
2	c/g	0-70	1389	23.0	9.4	16.2	17.7	160	36.5	−0.93
3	c/g	70-300	304	13.2	16.0	13.4	12.2	206	25.8	−0.68
4	c/g	300-700	133	13.5	5.7	11.9	7.1	301	26.4	−0.99
5	c/g	0-700	1863	21.1	9.4	16.1	15.1	196	33.9	−0.89
6	c/g	0-700$^{(*)}$	1863	23.6	−1.5	16.2	–	–	–	–
7	c/f	0-700	731	24.2	16.8	19.9	22.3	110	32.6	−0.65

Catalogs' abbreviations are: 'c' – for the Harvard CMT catalog, 'g' – for the GS-MT catalog, and 'f' – for the GS-FM catalog. 'n' – the number of matched pairs in each of two catalogs. Parameters a_0, σ are measured in degrees, a_1 is measured in degrees per unit of $|\Gamma|$, ϵ, and year, respectively. $^{(*)}$ the $|\Gamma|$-value from the GS-MT catalog is used. We match $m \geq 6$ CMT earthquakes with GS-MT and GS-FM events.

TABLE 7. Dependence of 3-D rotation pole location on focal mechanism and depth of earthquakes

#	Depth	FM	n	Φ	σ_Φ	η	σ_η	ψ	σ_ψ
1	0-35	Thr	681	25.08	15.58	89.88	40.27	192.6	105.0
2	0-35	S-S	480	22.78	15.32	89.19	29.30	199.7	97.7
3	0-35	Nor	294	30.13	18.48	88.15	38.50	186.4	100.9
4	0-35	All	1455	25.34	16.32	89.30	36.61	193.7	101.9
5	0-70	Thr	1142	24.80	15.71	89.62	42.48	192.3	104.6
6	0-70	S-S	553	22.89	15.27	89.14	30.42	198.8	99.1
7	0-70	Nor	392	29.90	18.45	87.46	37.75	186.3	102.5
8	0-70	All	2087	25.25	16.31	89.09	38.73	192.9	102.8
9	70-300	Thr	264	18.10	14.60	91.88	43.94	168.6	107.5
10	70-300	S-S	67	18.75	14.48	86.83	37.47	198.0	100.9
11	70-300	Nor	214	21.61	15.24	88.69	40.20	175.1	103.0
12	70-300	All	545	19.56	14.91	90.01	41.72	174.8	105.2
13	300-700	Thr	42	18.35	13.90	97.87	39.78	164.0	92.7
14	300-700	S-S	25	11.76	8.57	88.16	41.33	195.6	107.8
15	300-700	Nor	121	15.32	12.19	85.86	45.01	180.6	110.4
16	300-700	All	188	15.53	12.29	88.85	43.48	178.9	106.2
17	*EV*					90.0	39.17	180.0	103.9

We match $m \geq 5.3$ CMT earthquakes with GS-MT events. Focal mechanism (FM) is selected by using the CMT catalog: 'Thr' – thrust FM; 'S-S' – strike-slip FM; 'Nor' – normal FM. *EV* are expected values of parameters, if the rotation pole is distributed with uniform randomness on a reference sphere.

Fig. 1: USGS 1980/1/1−−2000/7/31, shallow: − n56; −. n58; −− n60; − n62

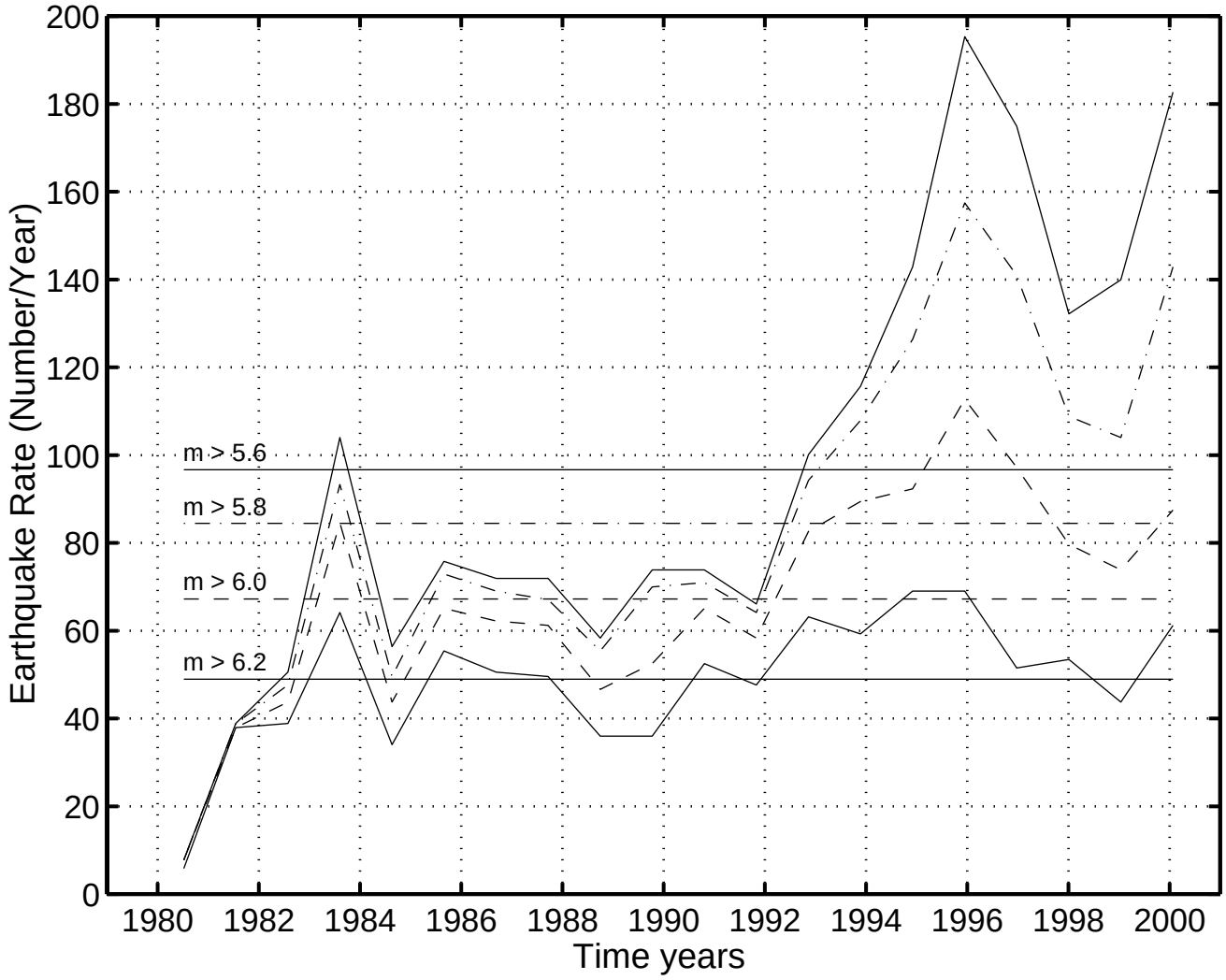


Fig. 2: Harvard catalog 1977/1/1−−2000/12/31: m>5.8, : G−R

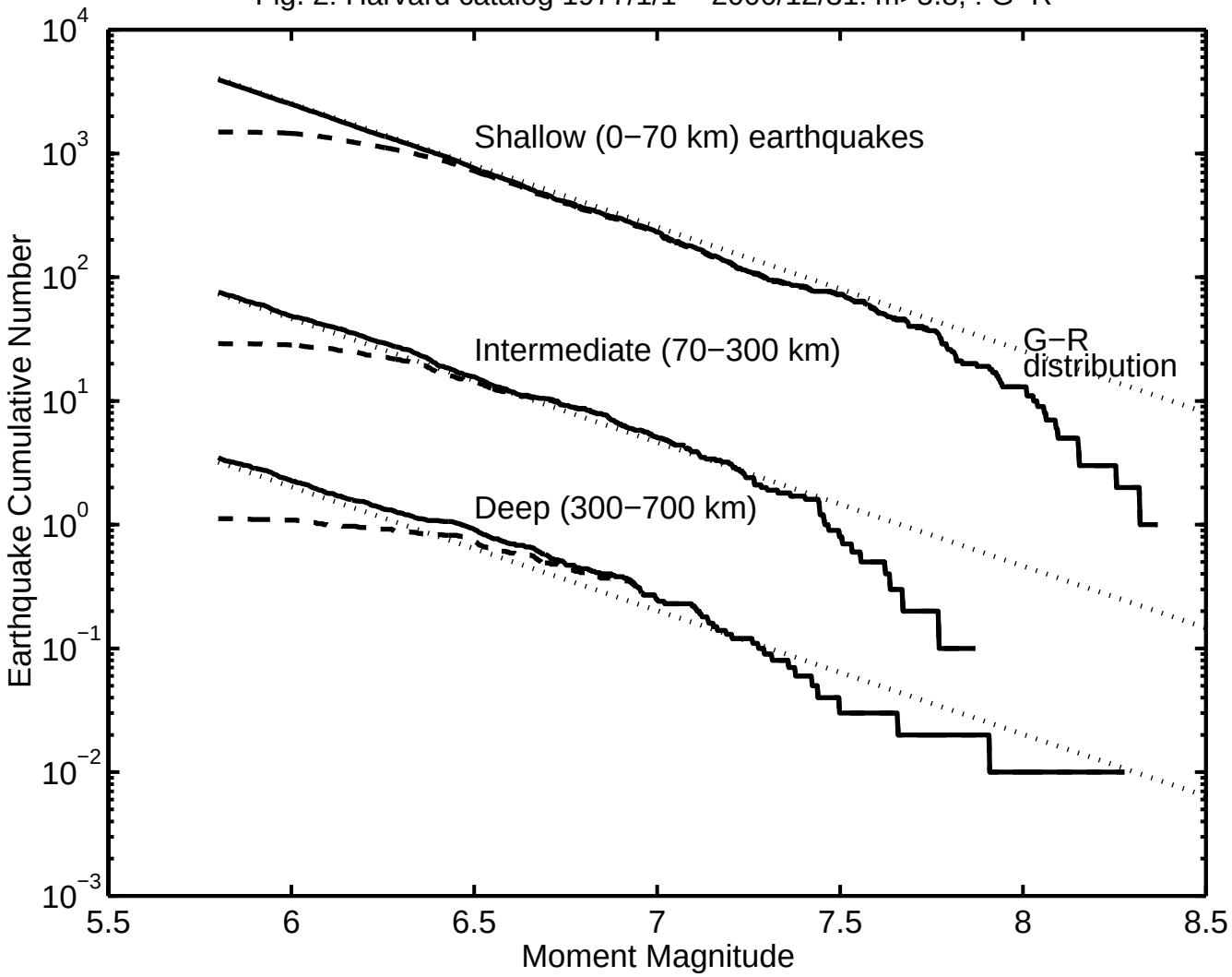


Fig. 3: CMT 1977−2000, β (M_v): − 0−70km; −− 70−300km; : 300−700km

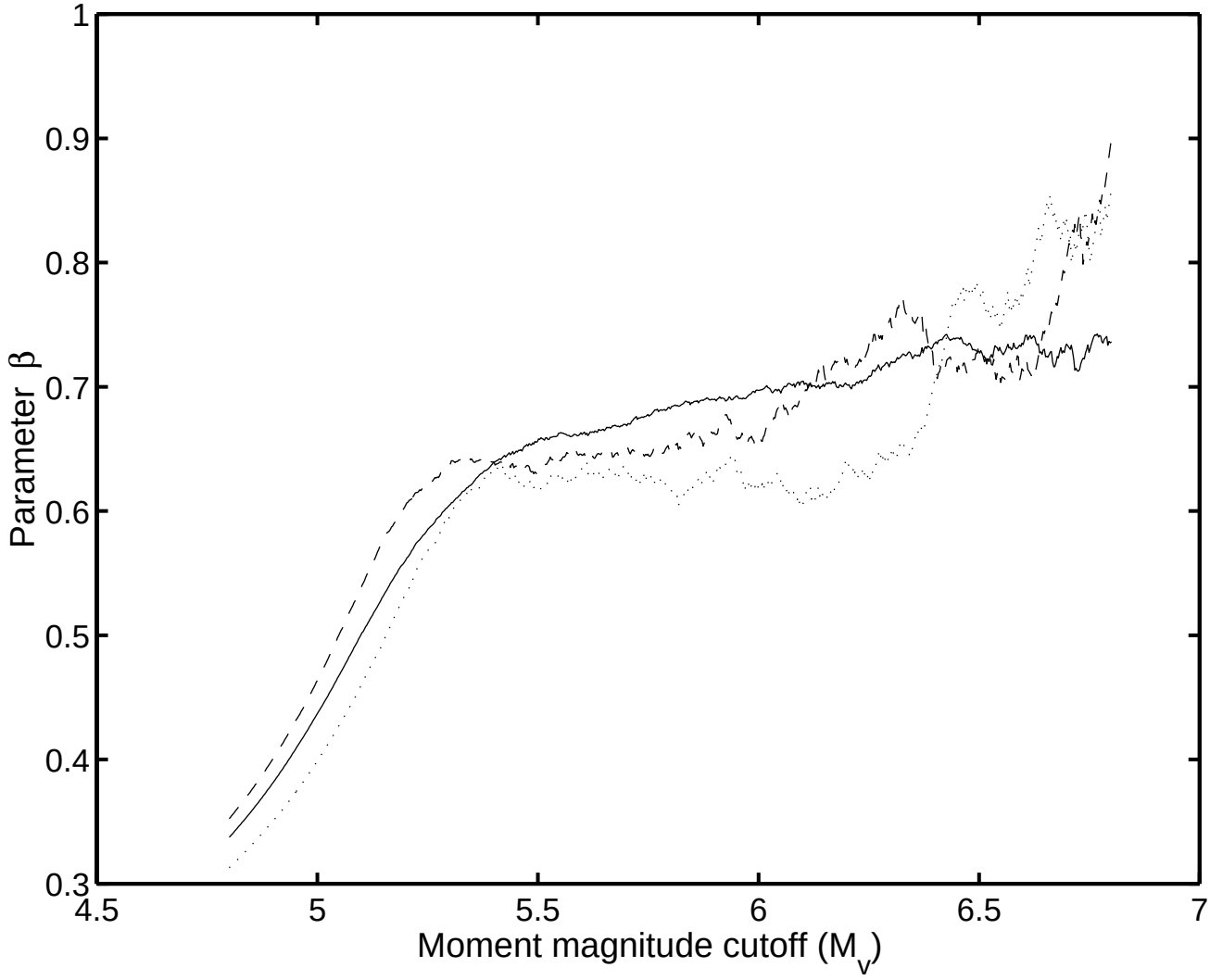


Fig. 4: CMT 1977−2000, α: − 0−70km; −− 70−300km; : 300−700km

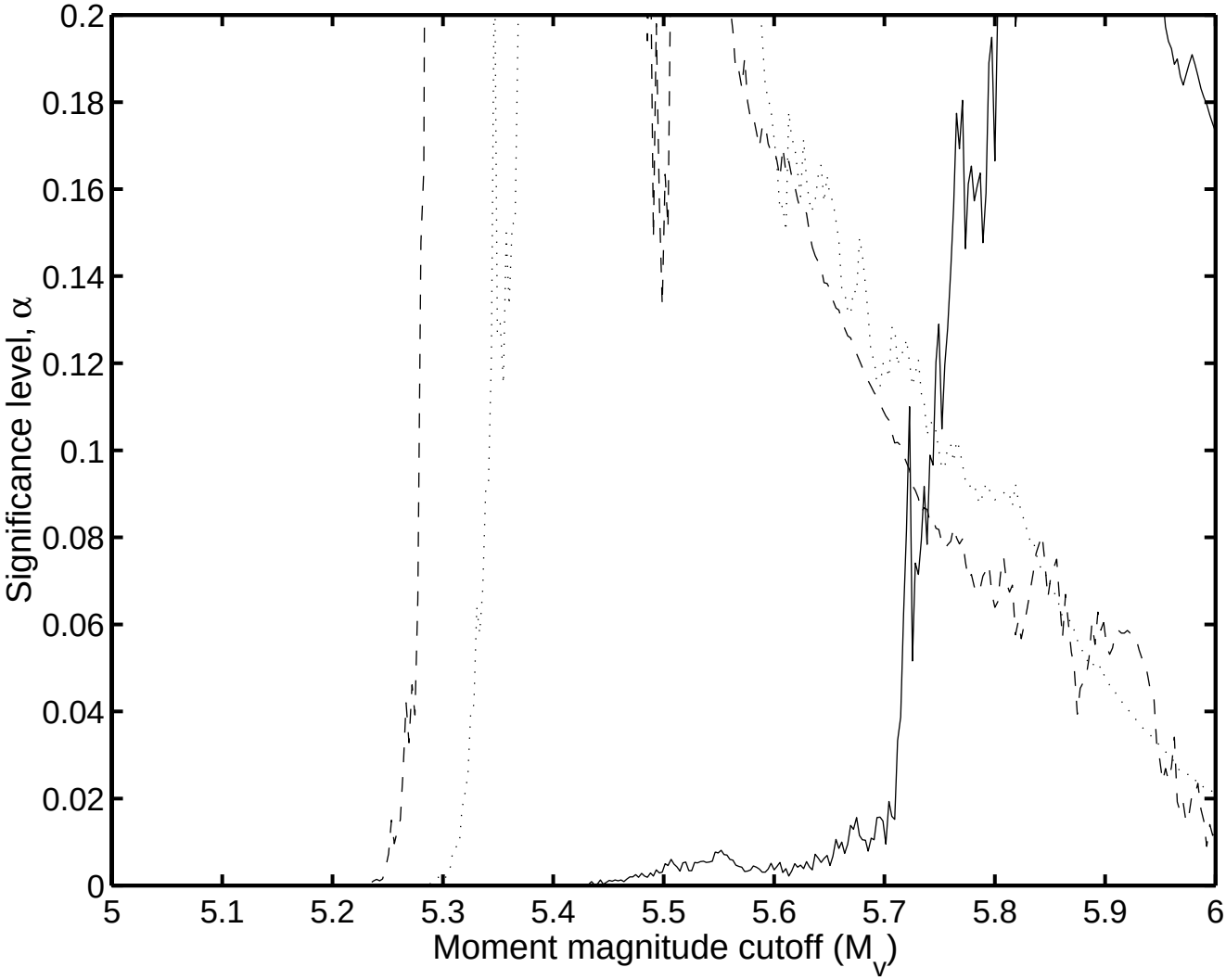


Fig. 5: USGS 1980−−2000/7/31 (M>5.0, 0−700km, unmatched CMT), Start−val = 1980

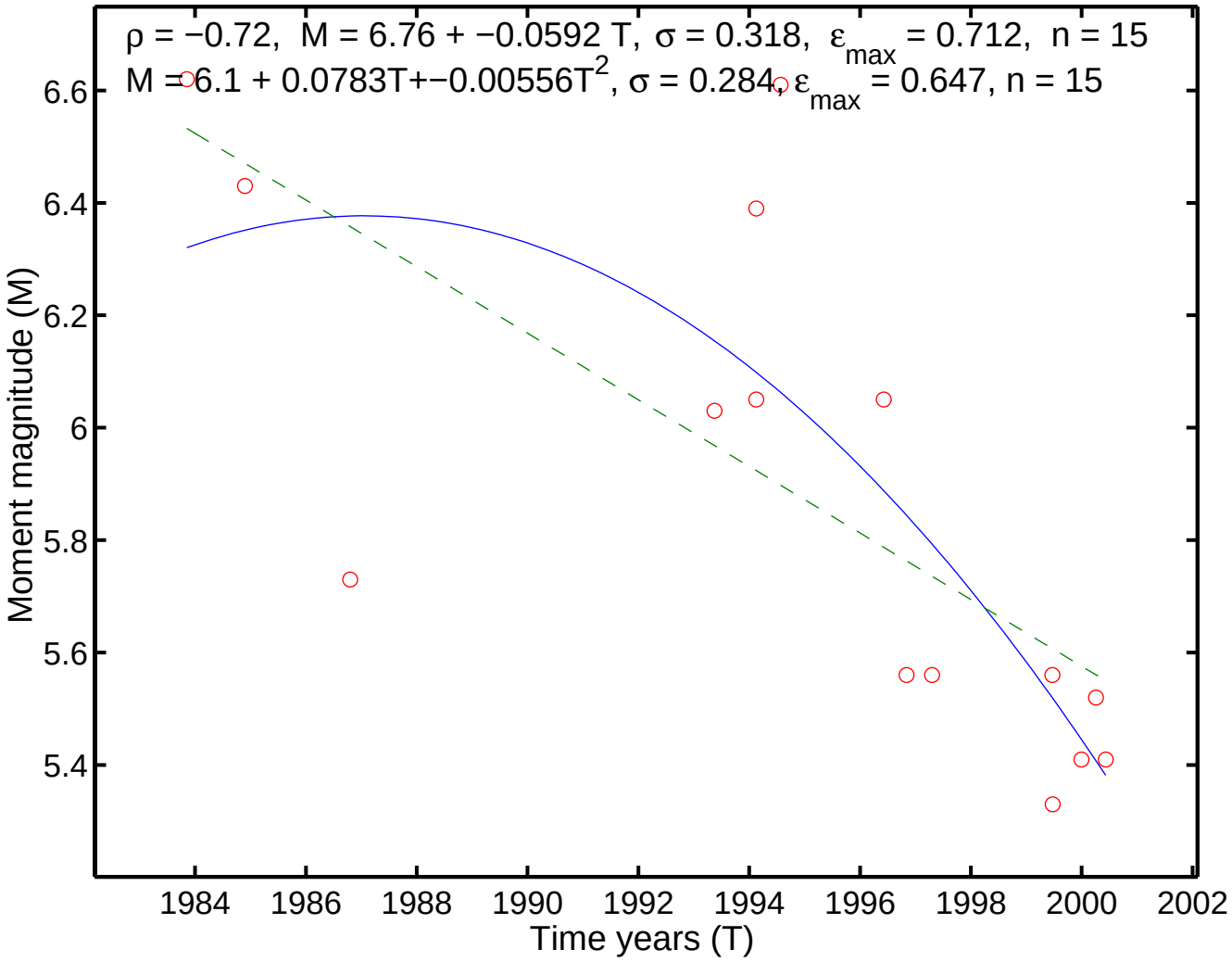


Fig. 6: Time next eq, 0−700km, 1977−2000: − 6>M>5.5, −− 6.5>M>6, −. 7>M>6.5, : 7.5>M>7

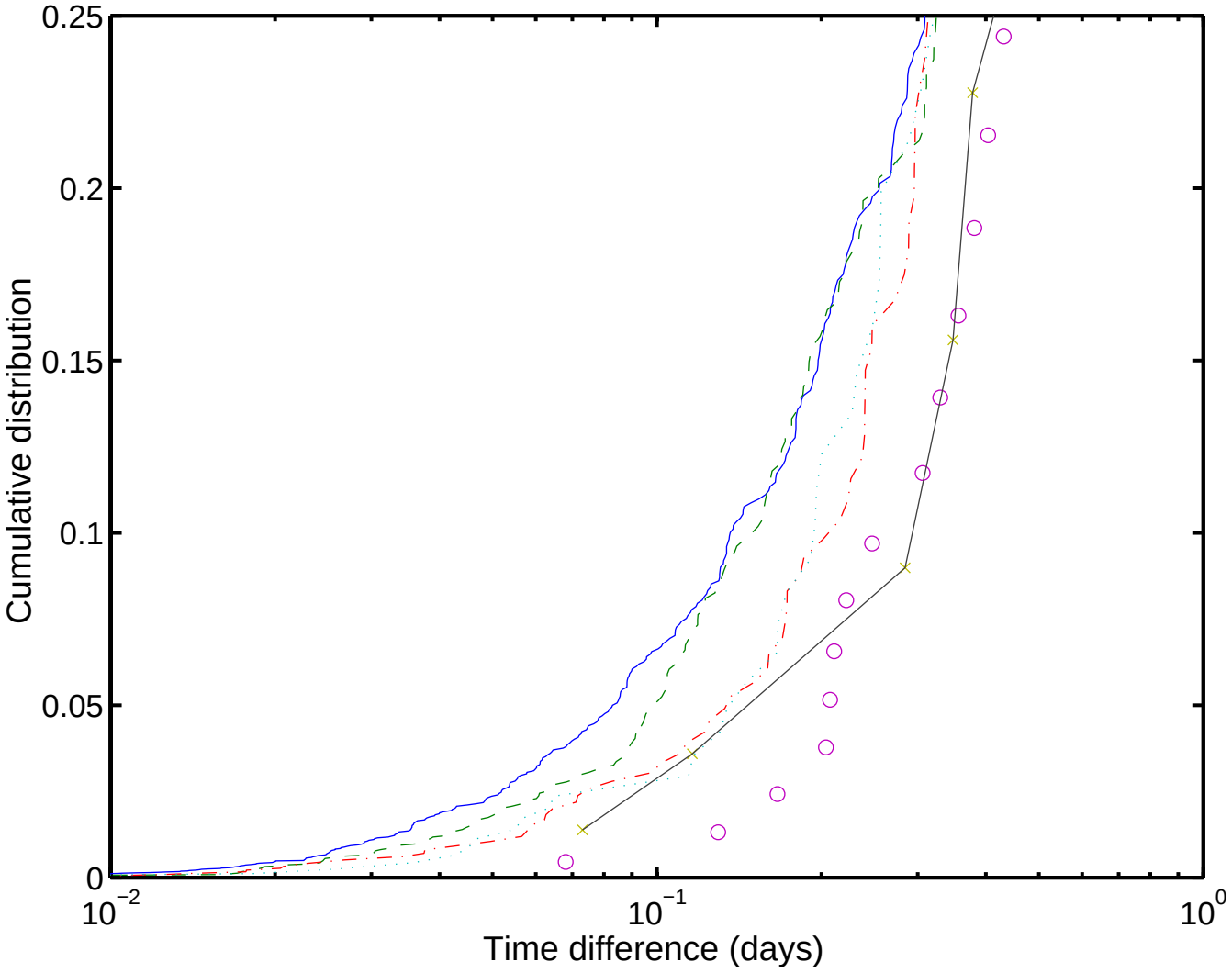


Fig. 7: CMT−PDE 1977−−2000 (M>5.3, R_{max}=250km, 0−70km), match, M_w/m_b, Start−val = 6

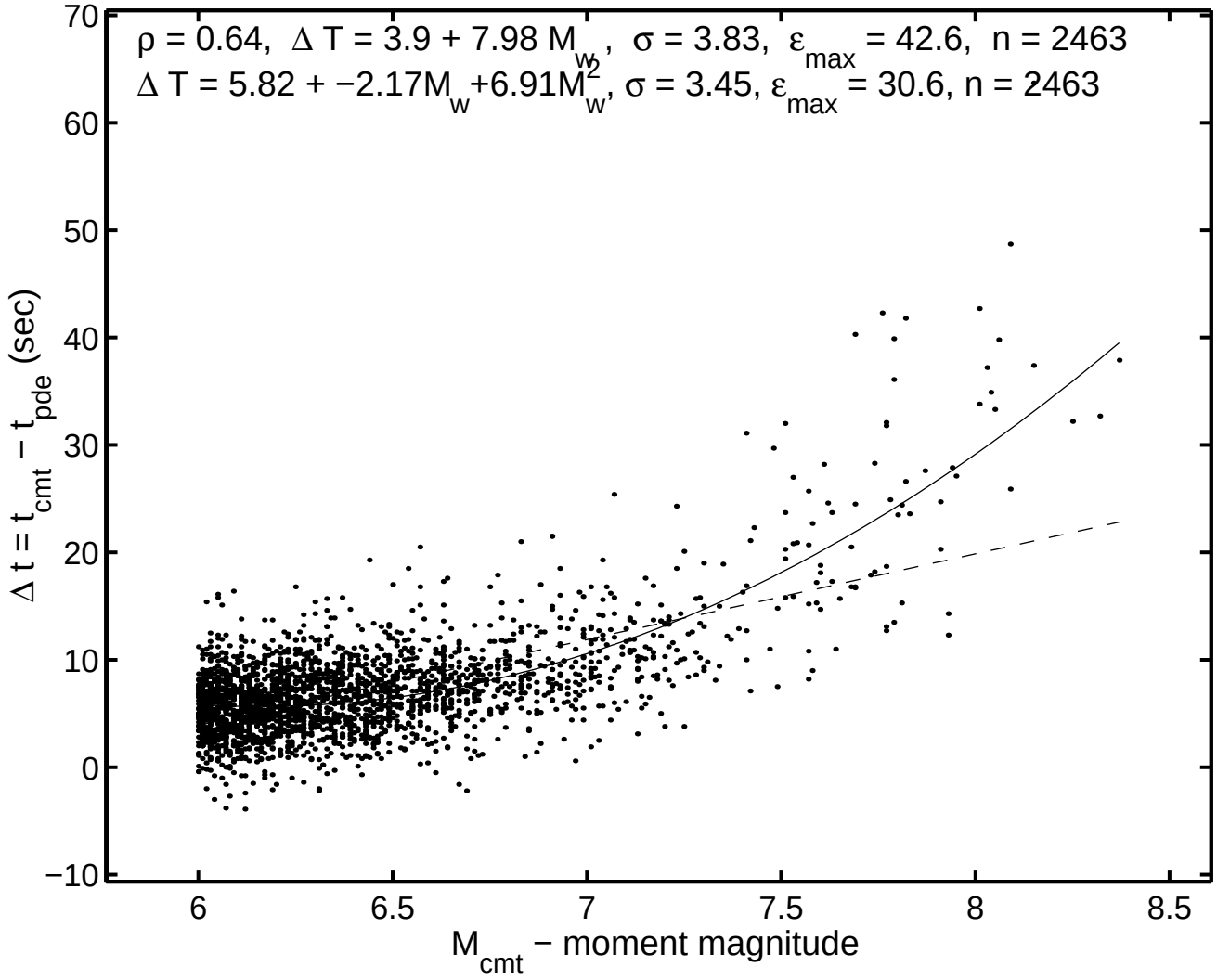


Fig. 8

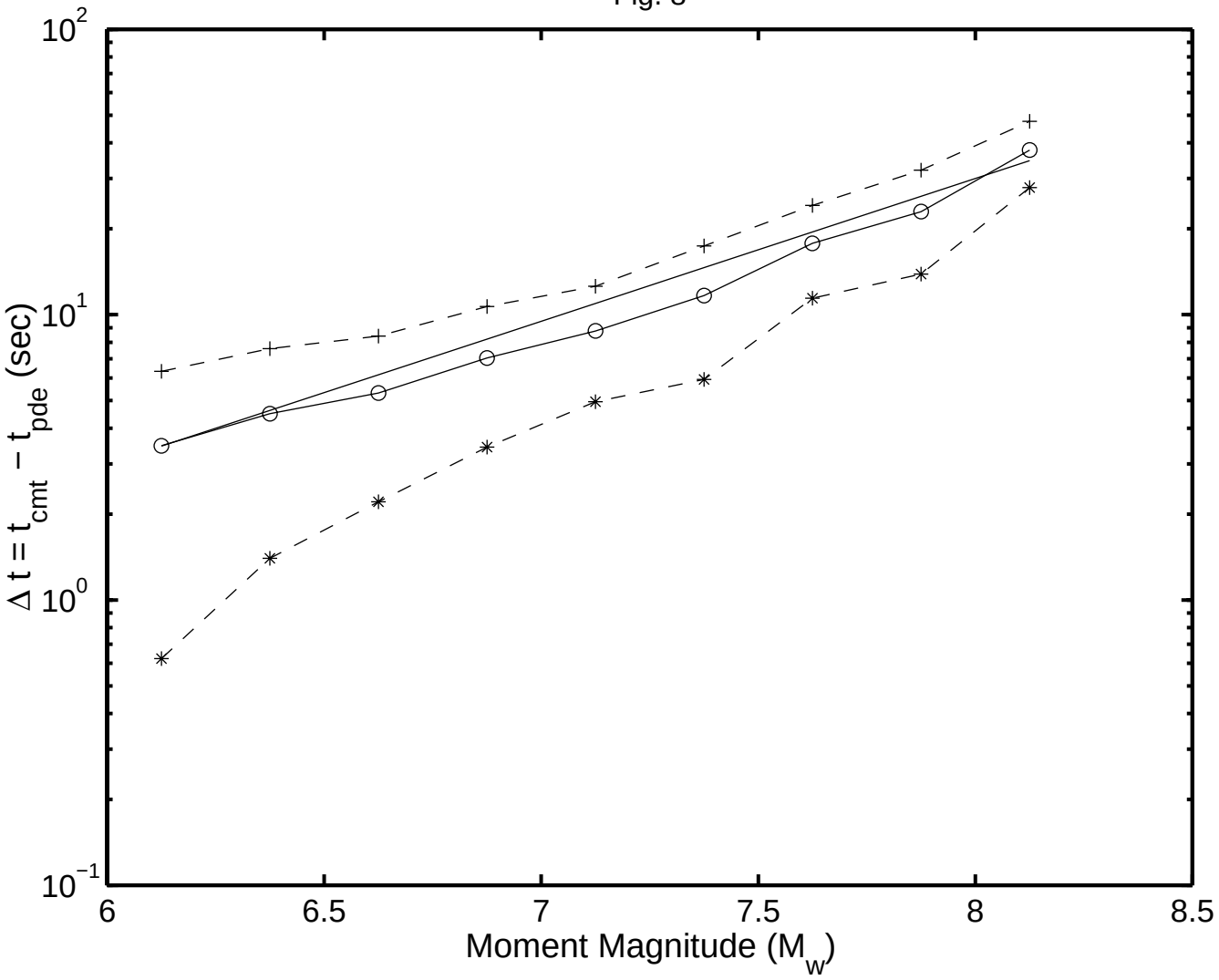

10^2
10^1
10^0
10^{-1}
$\Delta t = t_{cmt} - t_{pde}$ (sec)
6
6.5
7
7.5
8
8.5
Moment Magnitude (M_w)

Fig. 9: CMT/PDE, 1977−2000/12/31, 0−70km, M>=6.0, <R> (epicentral distance) +/− σ

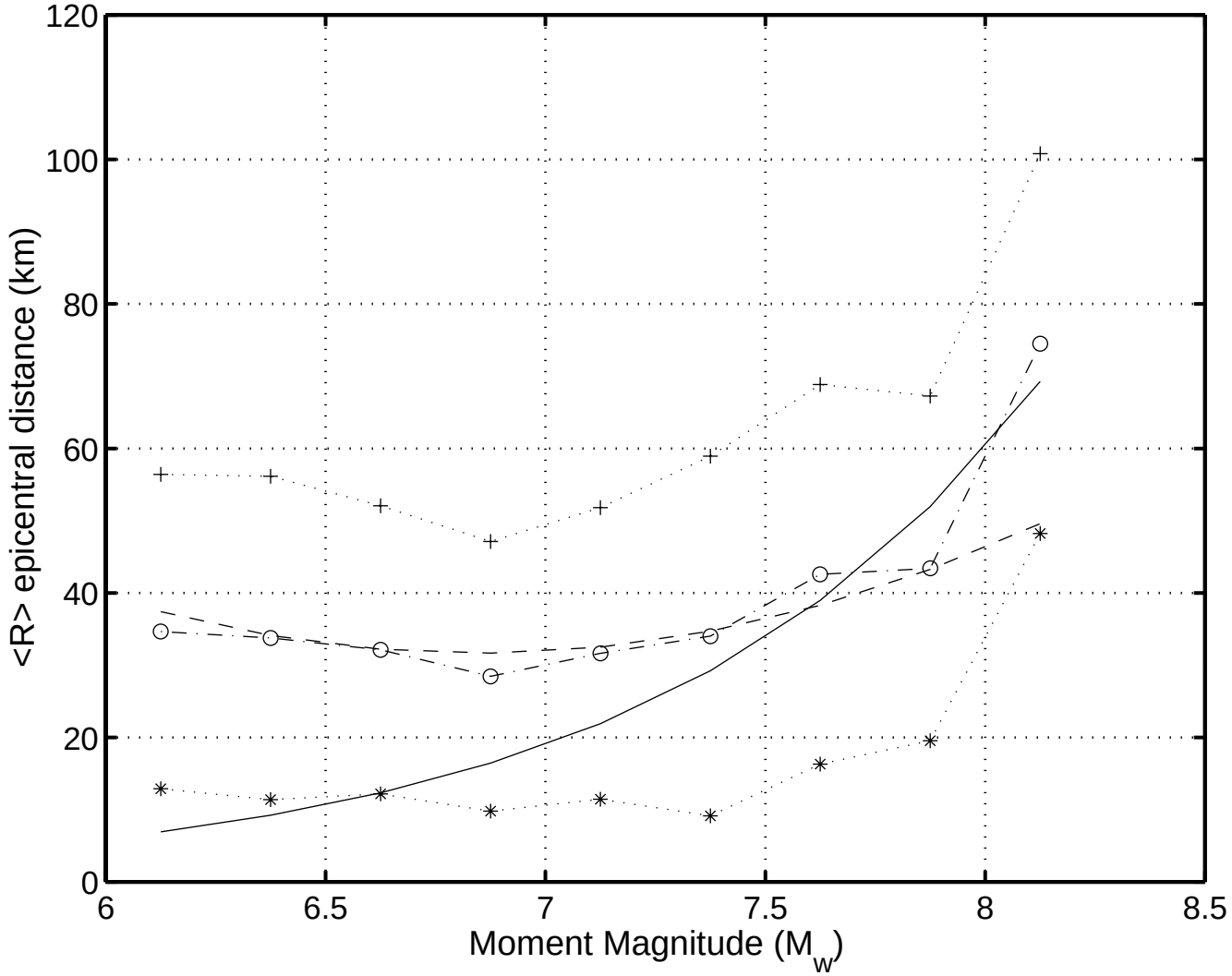


Fig. 10: PDE/ISC, R_{max}=250km, 1971−1998, 0−70km, M>=5.0, <R> (epicentral distance) +/− σ

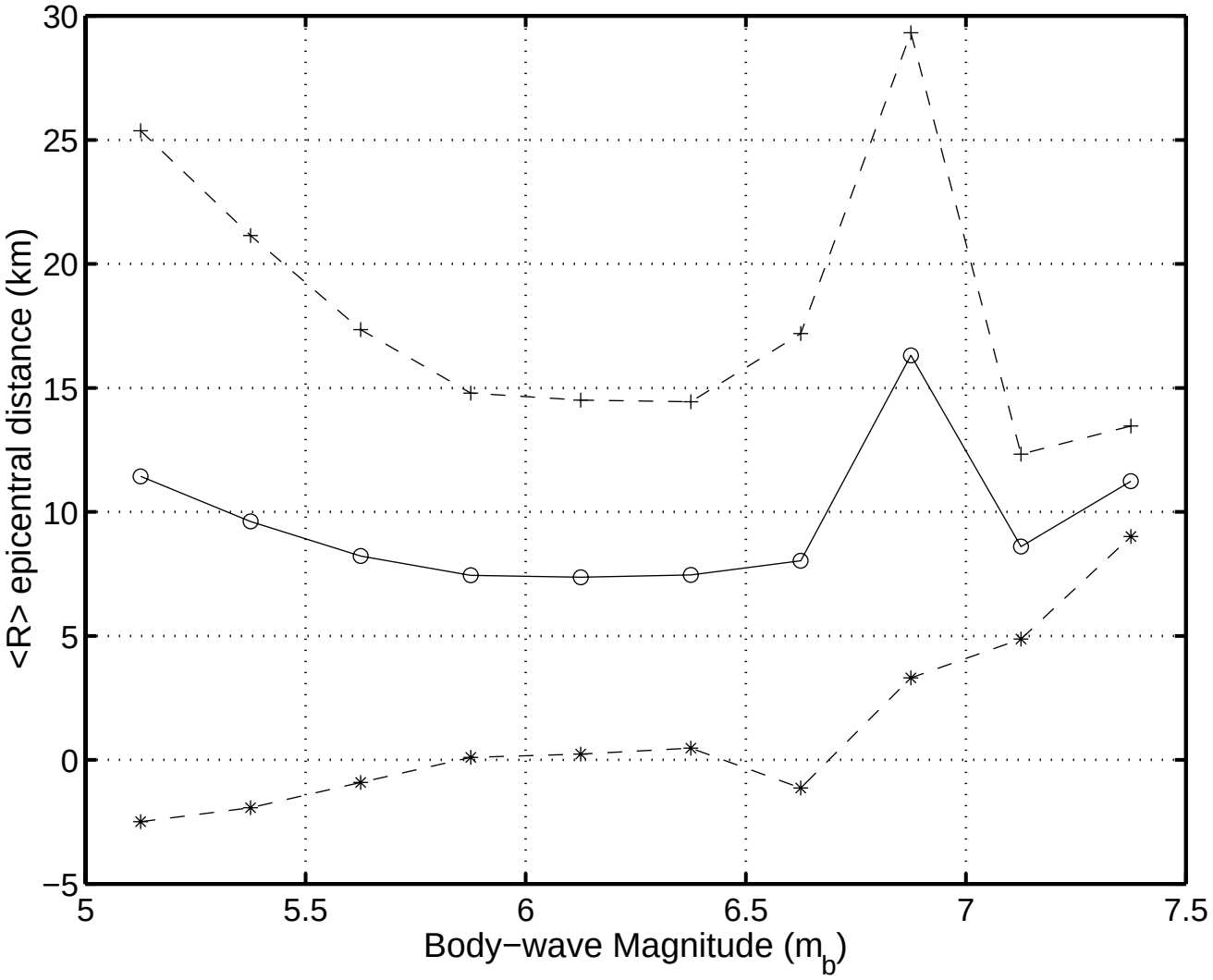


Fig. 11: PDE/ISC 1971/1/1−−1998/12/31 (m_b>5.0, 0−70km), matched

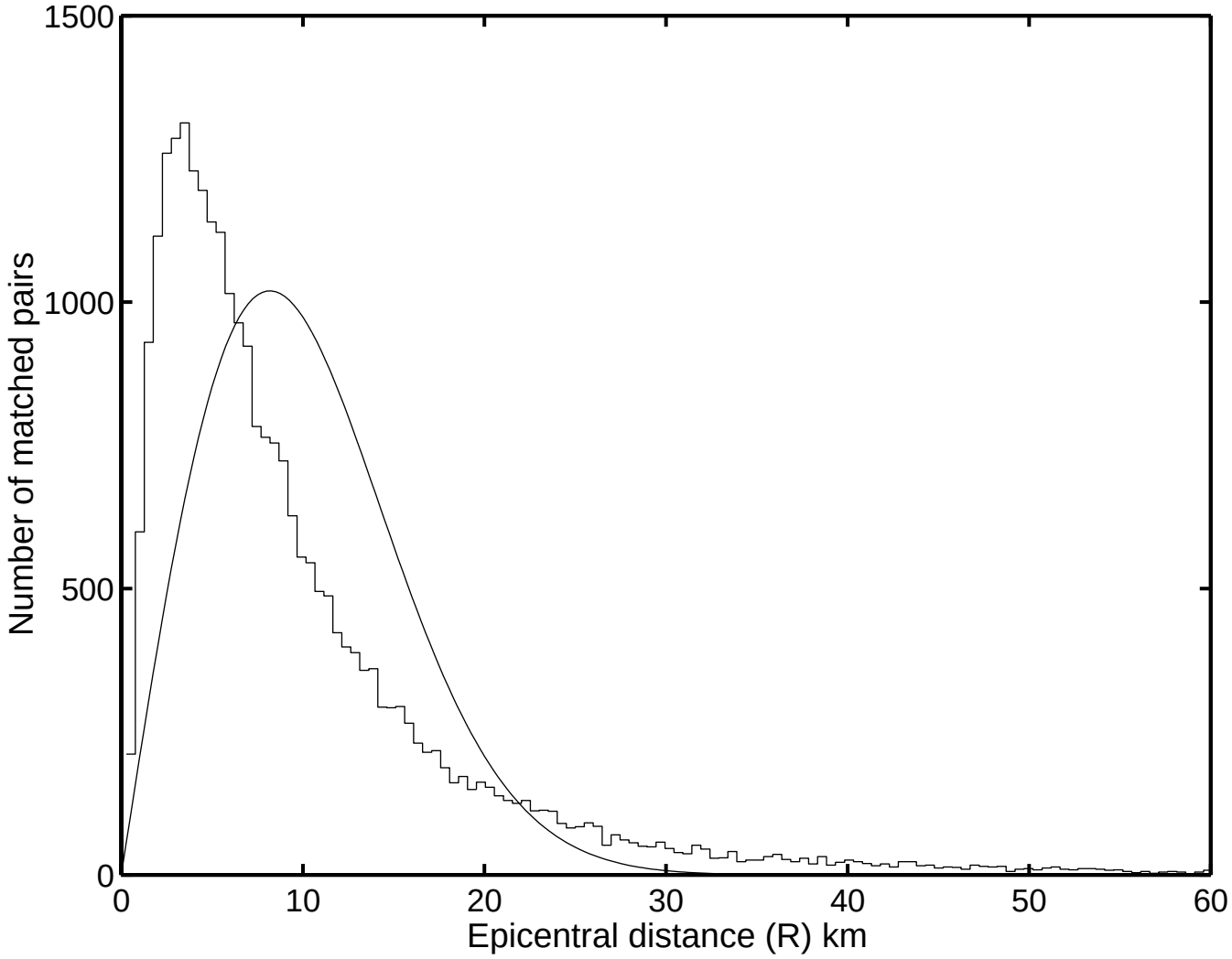


Fig. 12: CMT/USGS 1980−−2000/07/31 (M>6.0 pairs, shal. 0−35km, R<140km), Start−val = 6

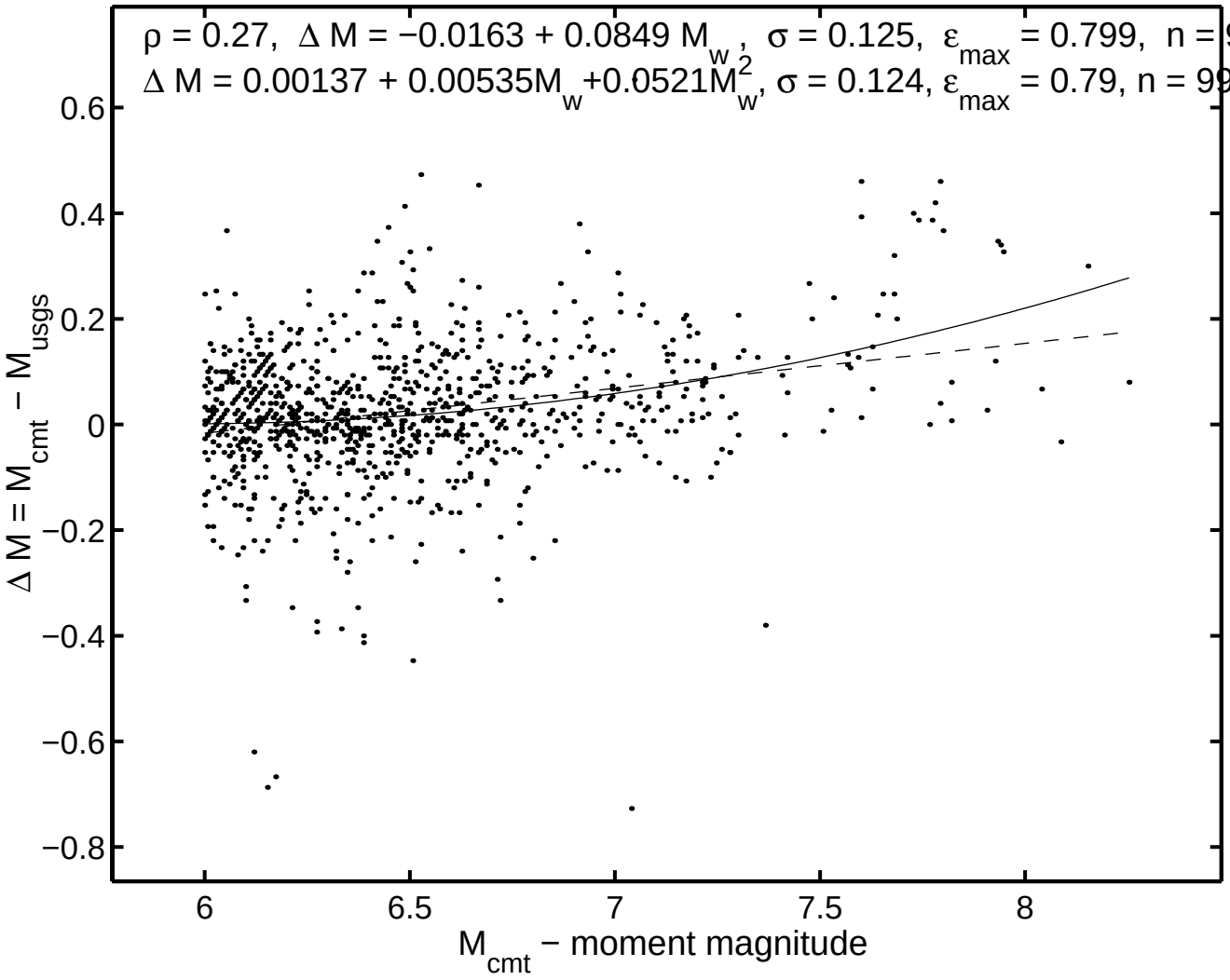


Fig. 13: CMT/PDE 1977−−2000 (M_w>6.0 pairs, matched, all 0−35km, R<250km), Start−val = 6

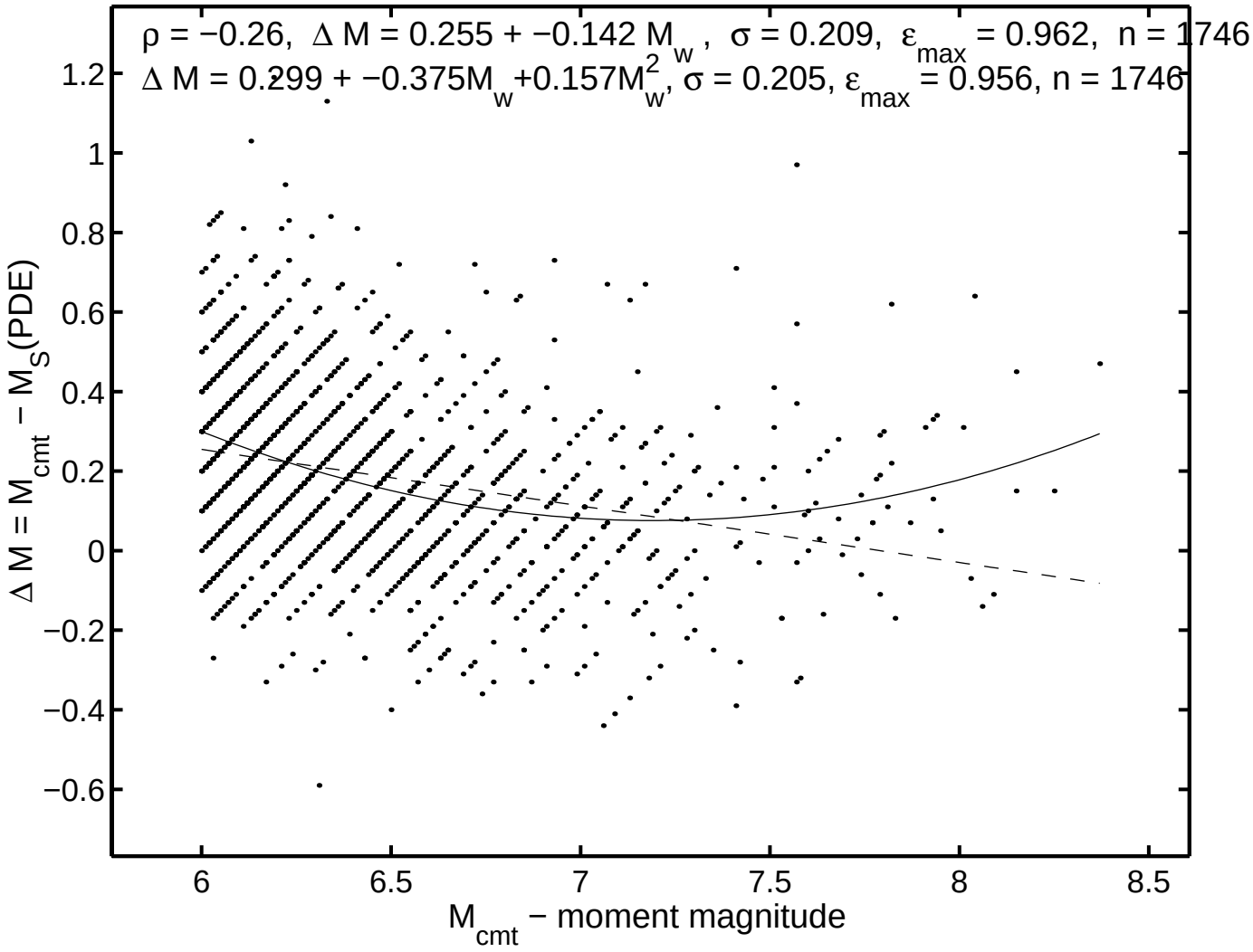


Fig. 14: CMT/PDE 1977−−2000 (M_w>5.3 pairs, matched, 0−70km, R<250km), Start−val = 5.3

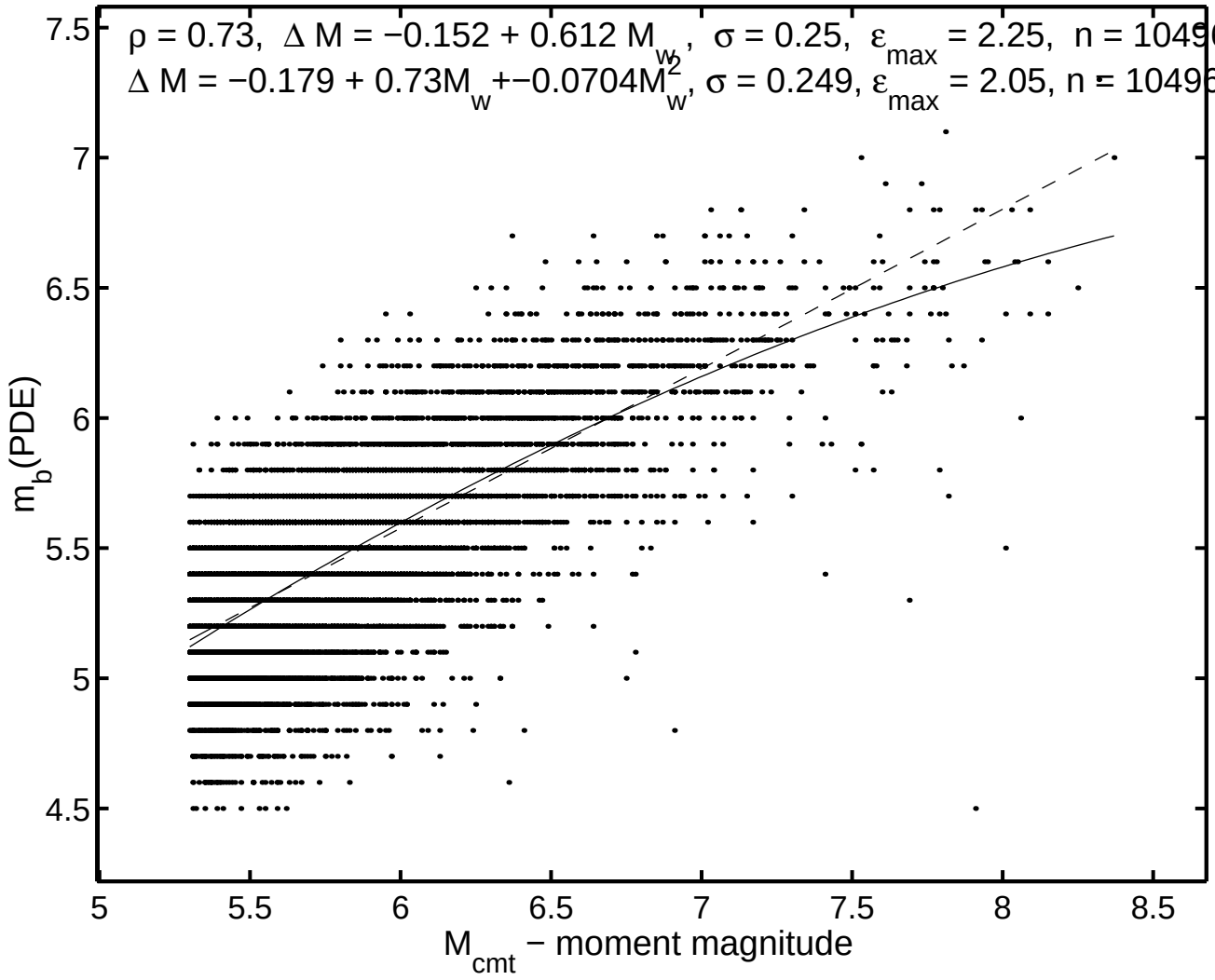


Fig. 15: CMT/USGS, R_{max}=140km, 1980−2000/7/31, 0−700km, M>=6.0, Δ M +/− σ

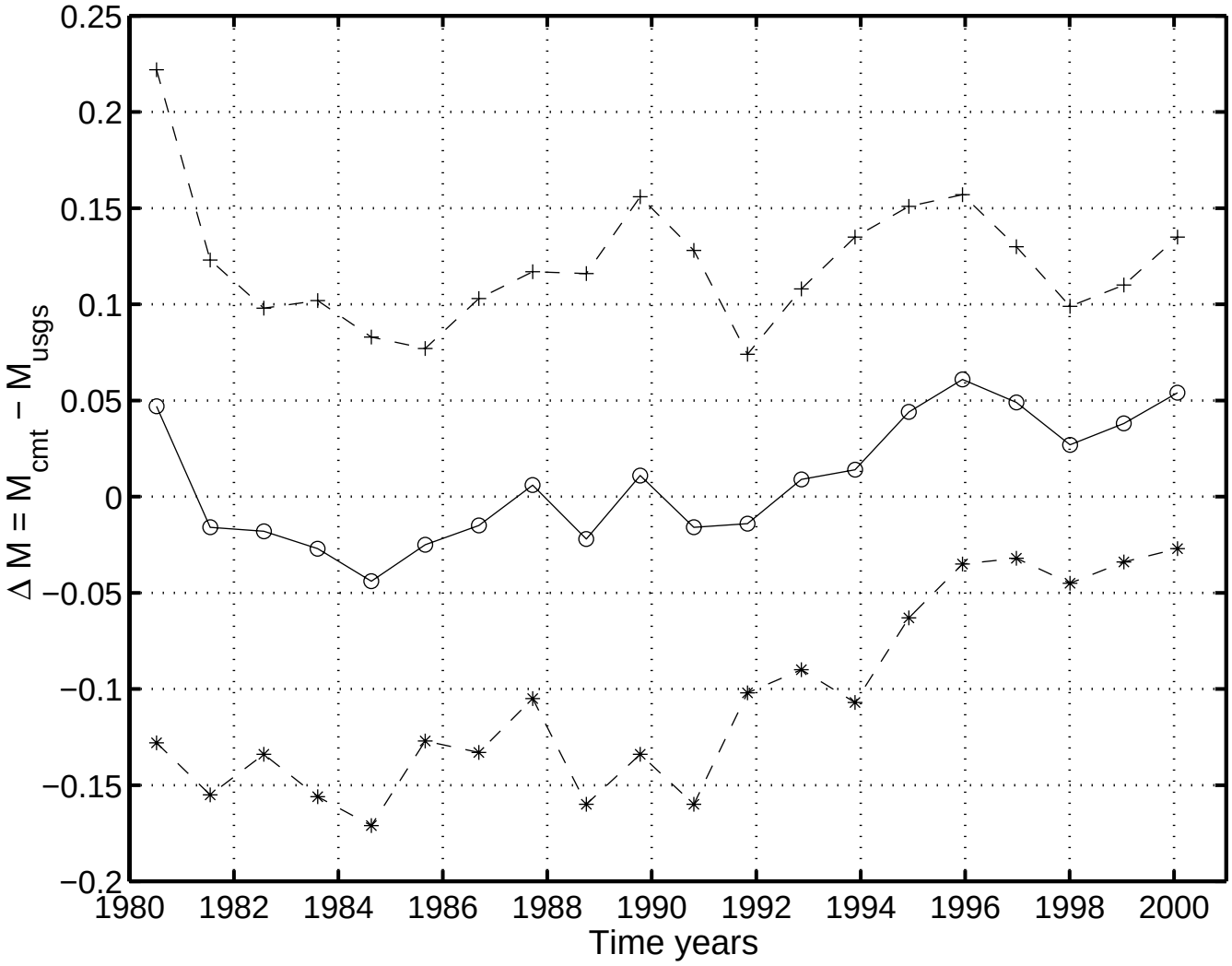


Fig. 16a: CMT 1977−−2000/12/31 (M>5.3 shallow, all, 0−70km, n_{comp}= 6) Relative error, ε_{rr}

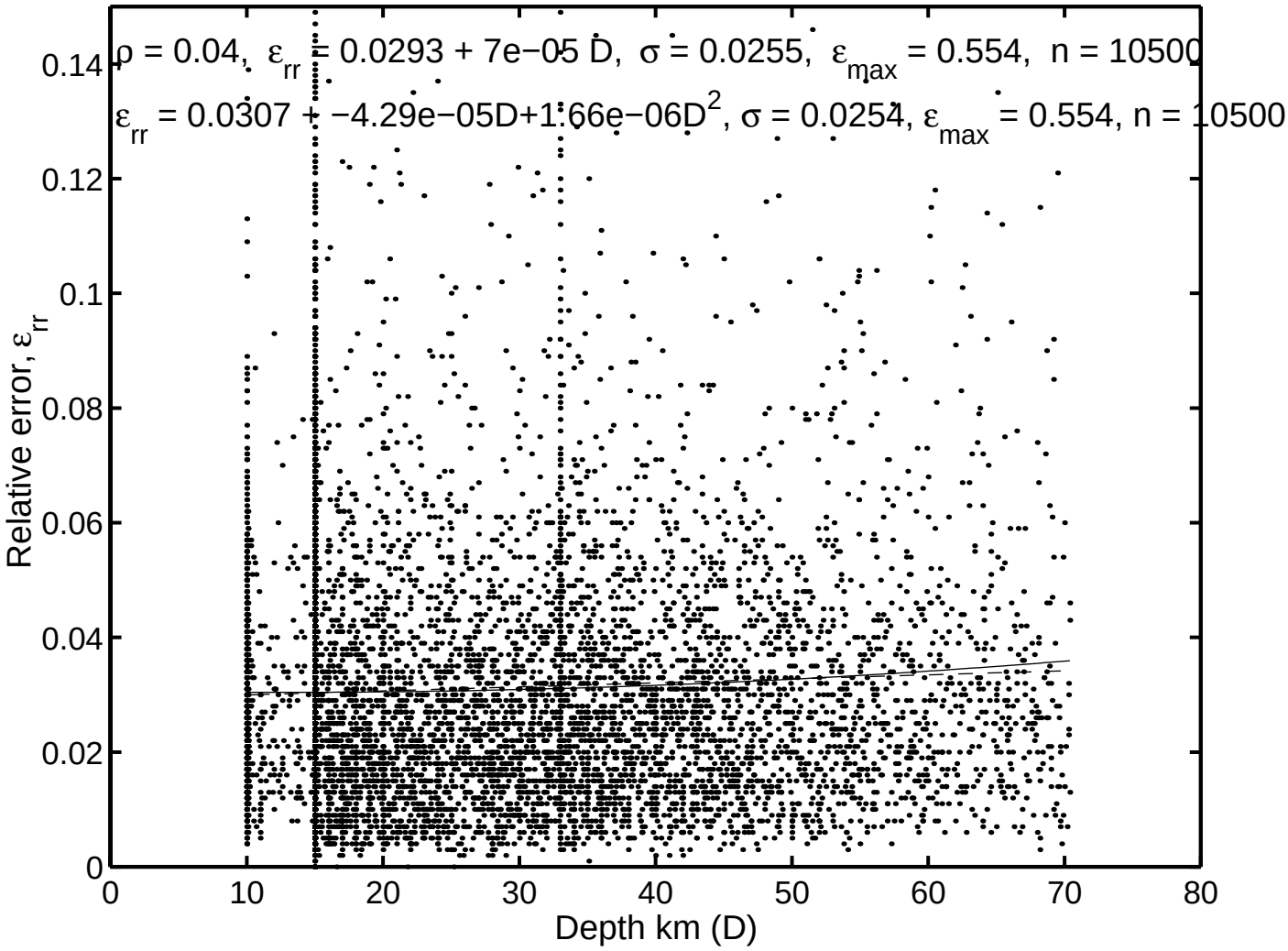


Fig. 16b: CMT 1977−−2000/12/31 (M>5.3 shallow, all, 0−70km, n_{comp} = 6) Relative error, $\varepsilon_{r\theta}$

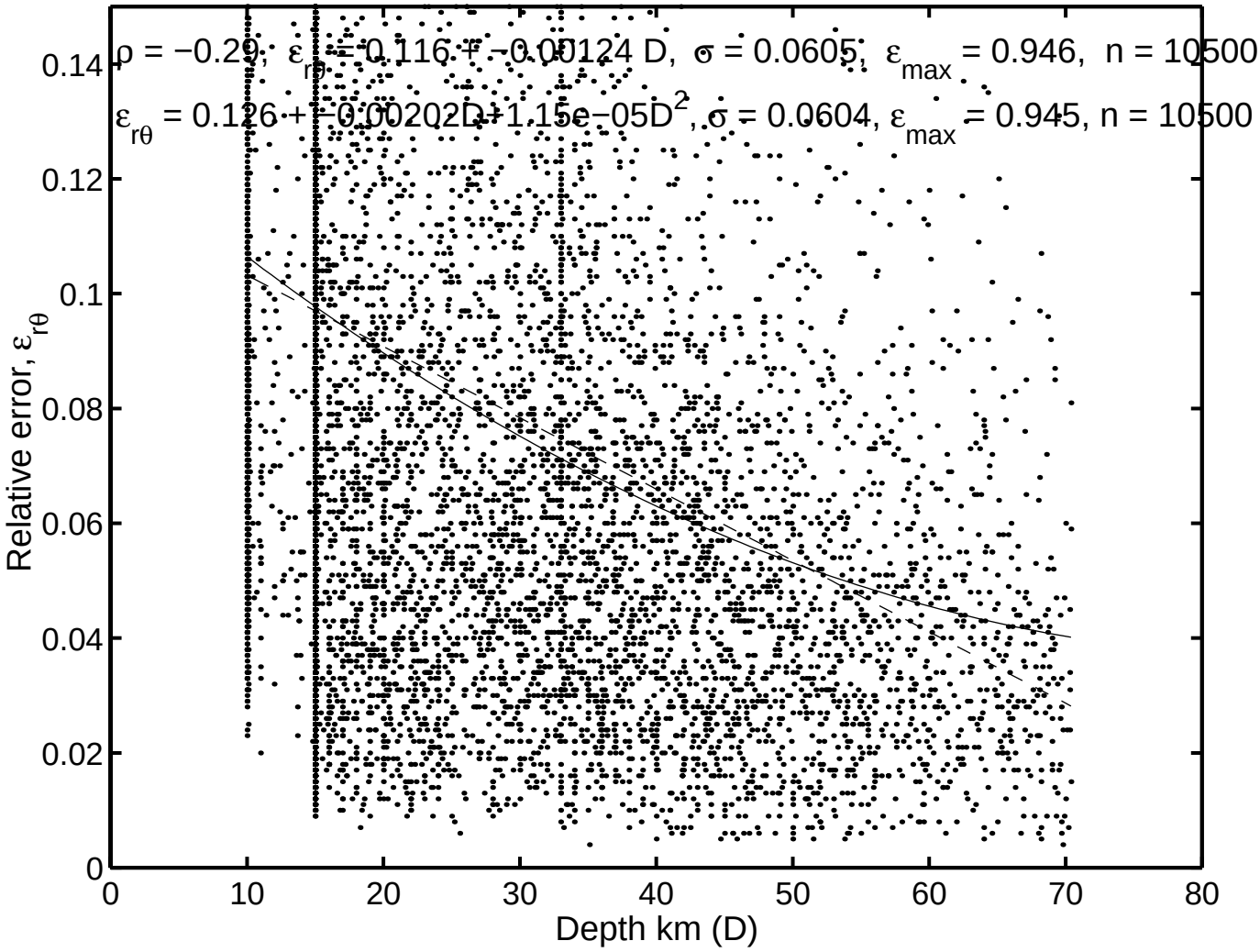


Fig. 17: CMT 1977−−2000 (M>5.3, 0−70km) CLVD Index,Γ

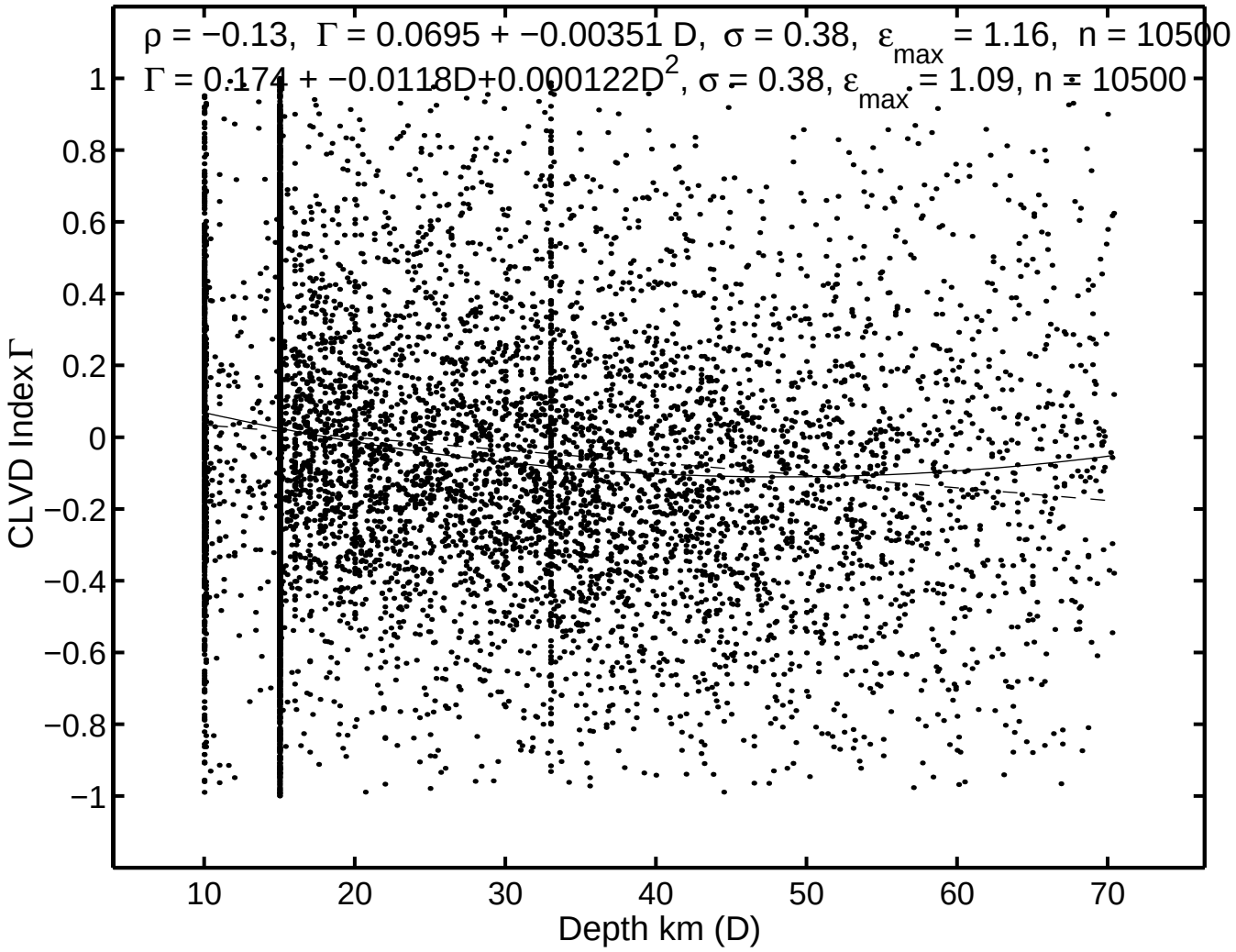


Fig. 18: CMT/USGS 1980−−2000/07/31 (M>5.3 pairs, shallow 0−70km, R<140km), Γ

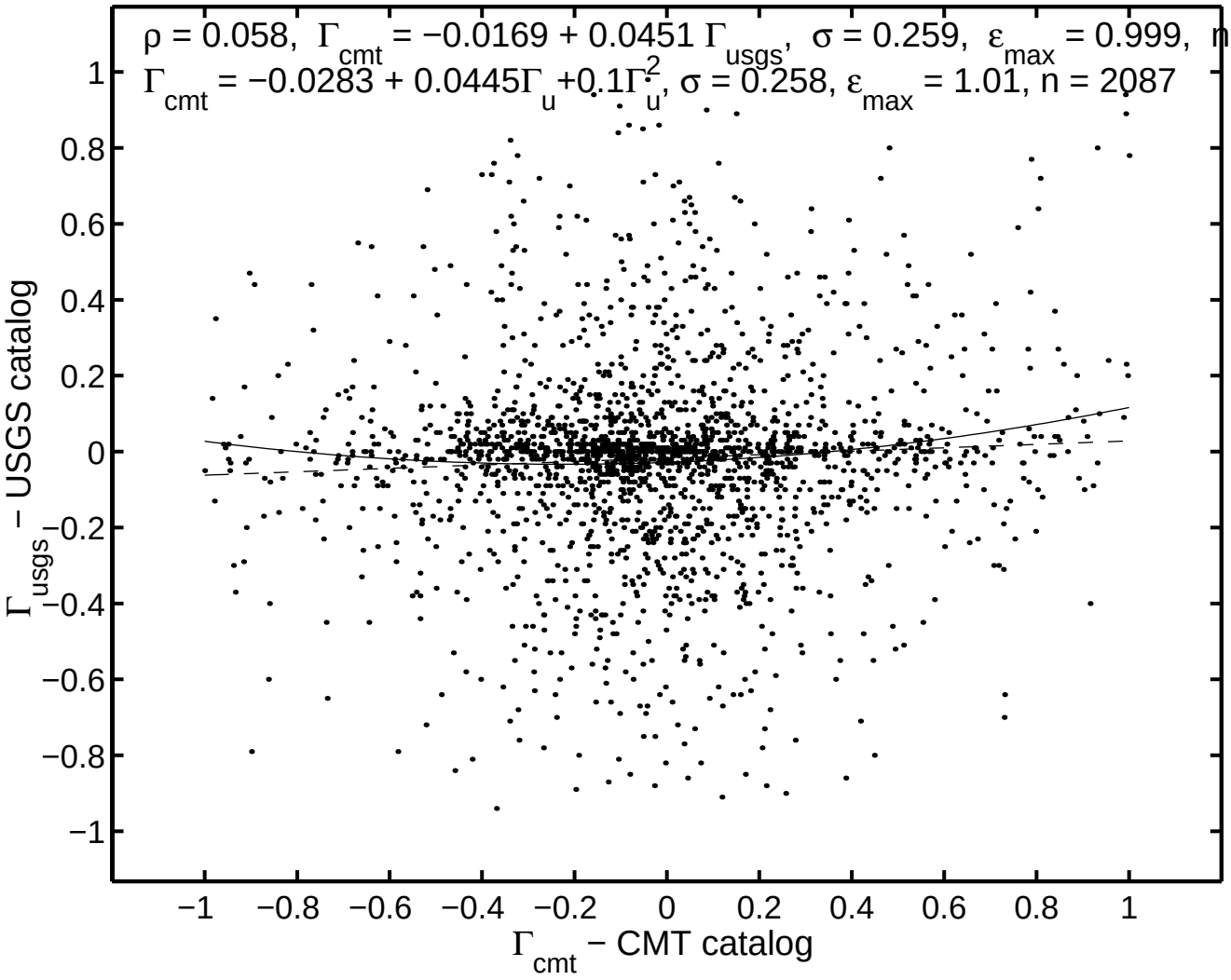


Fig. 19: CMT−USGS 1980−2000/07/31, Φ hist., m>=6, 0−70 km; − Rayleigh, −. Maxwell, −− Cauchy

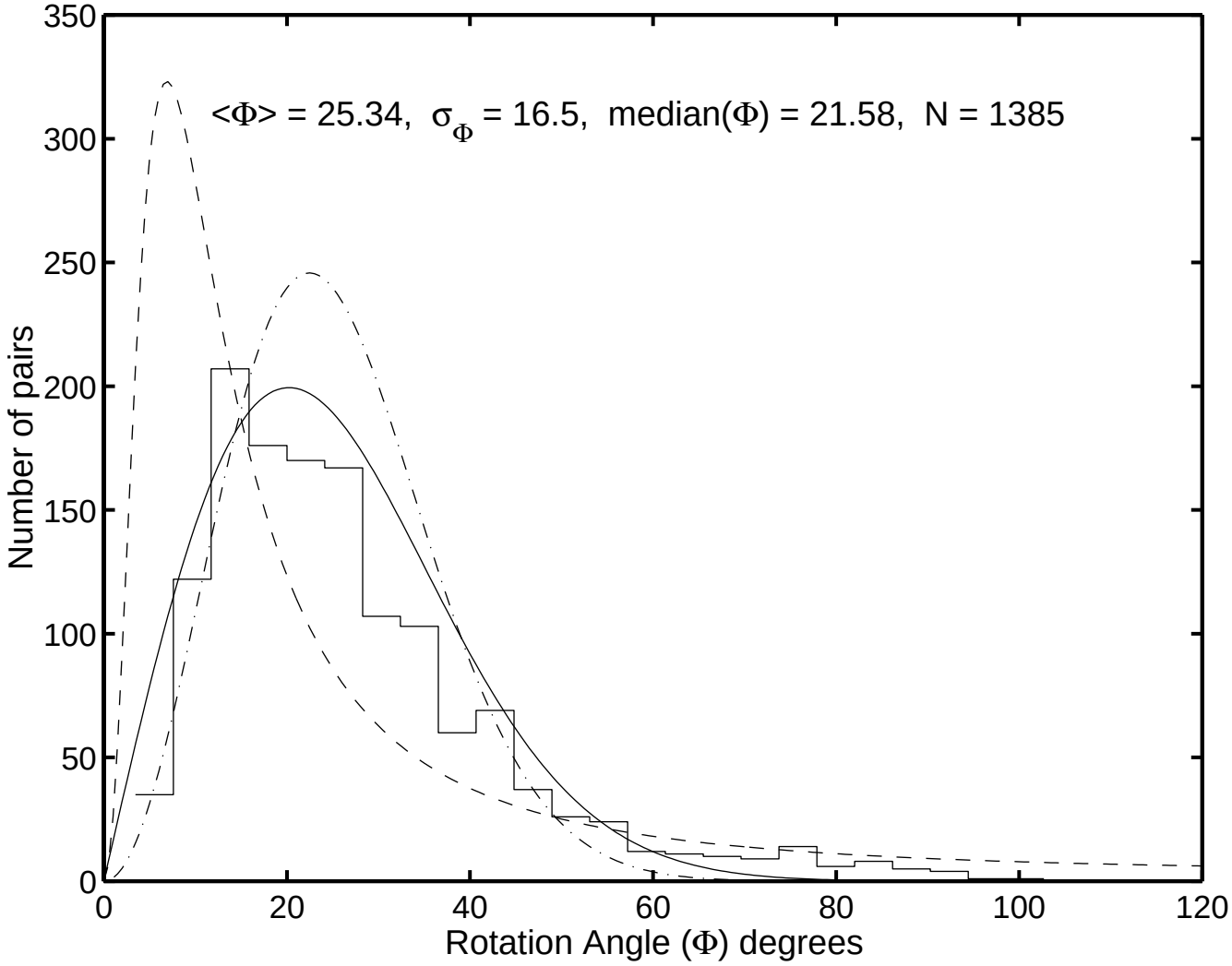


Fig. 20: CMT/USGS 1980−−2000/07/31 (M>6 pairs, all 0−700km, R<140km), Start−val = 1980

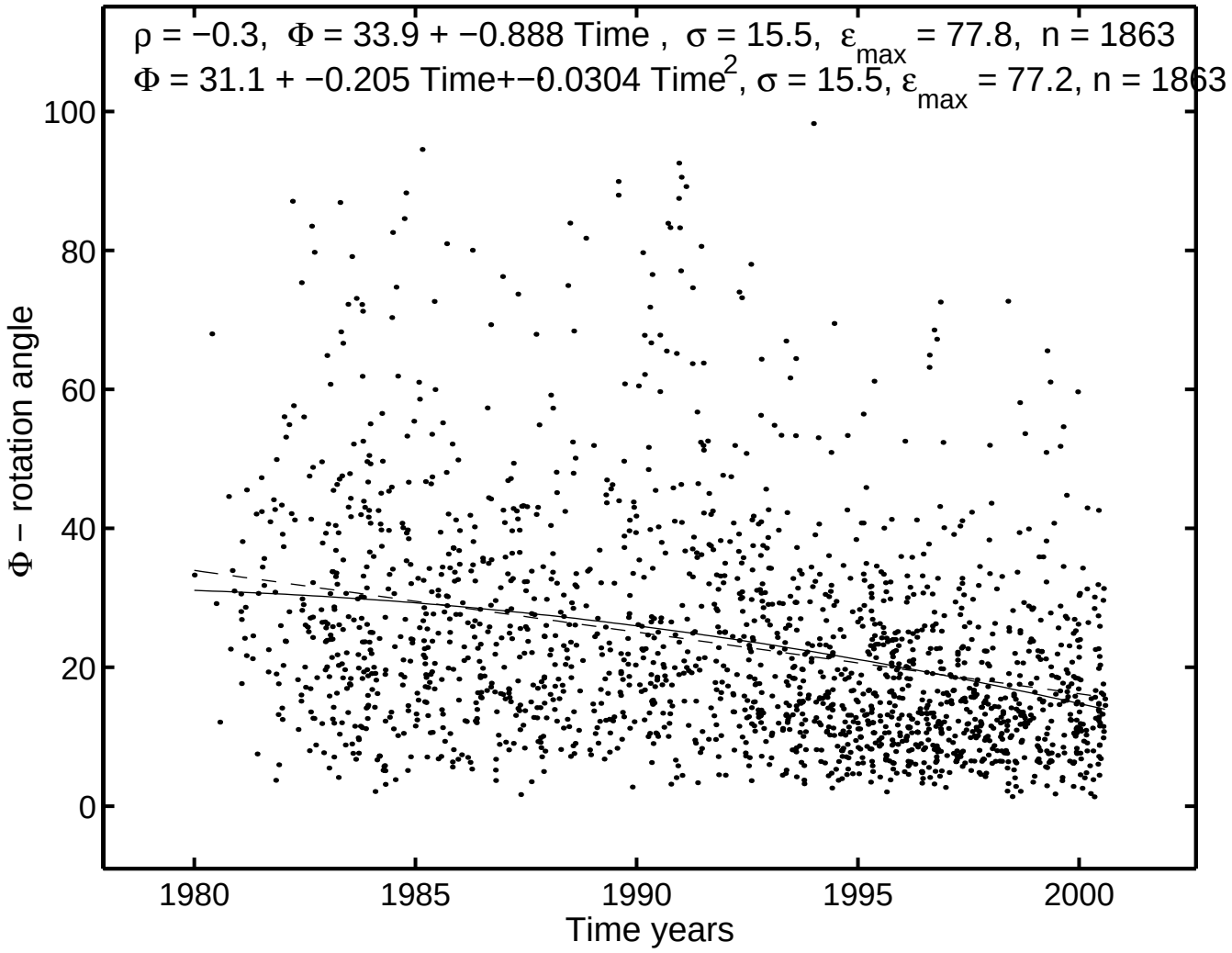


Fig. 21a: CMT−−GS−MT Rotation distr. hist., CMT−strike−slip, m>=5.3, 0−35 km

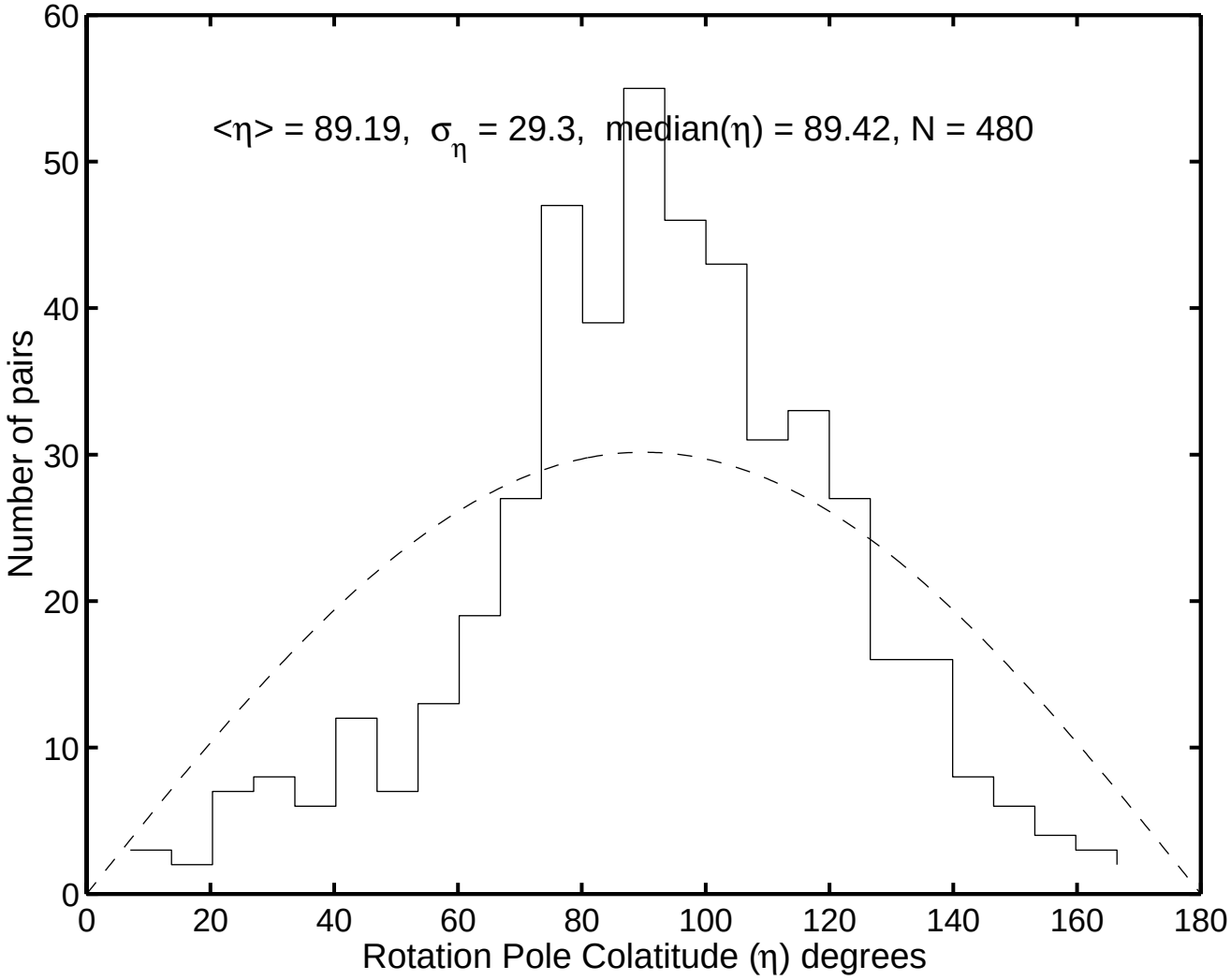
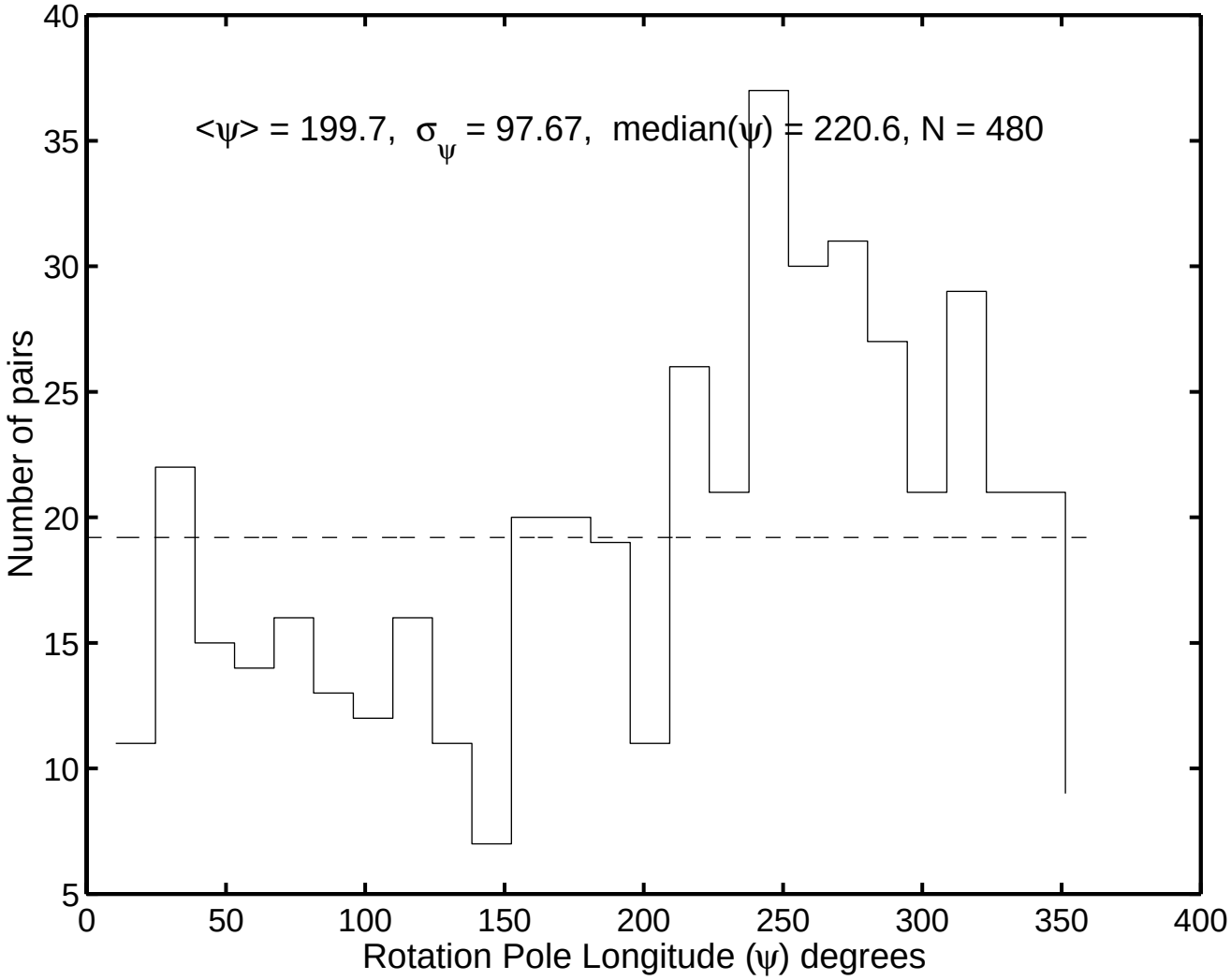
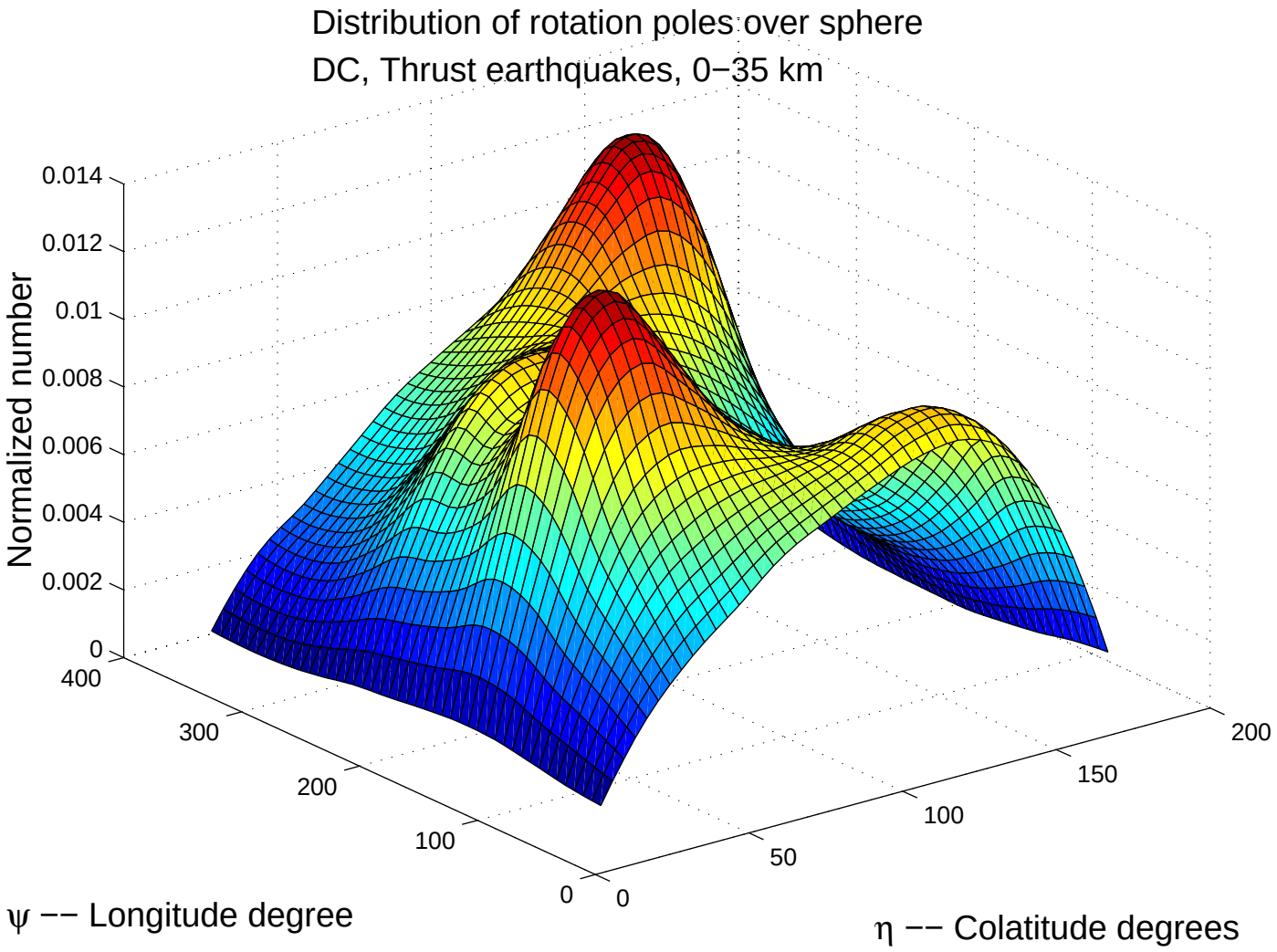
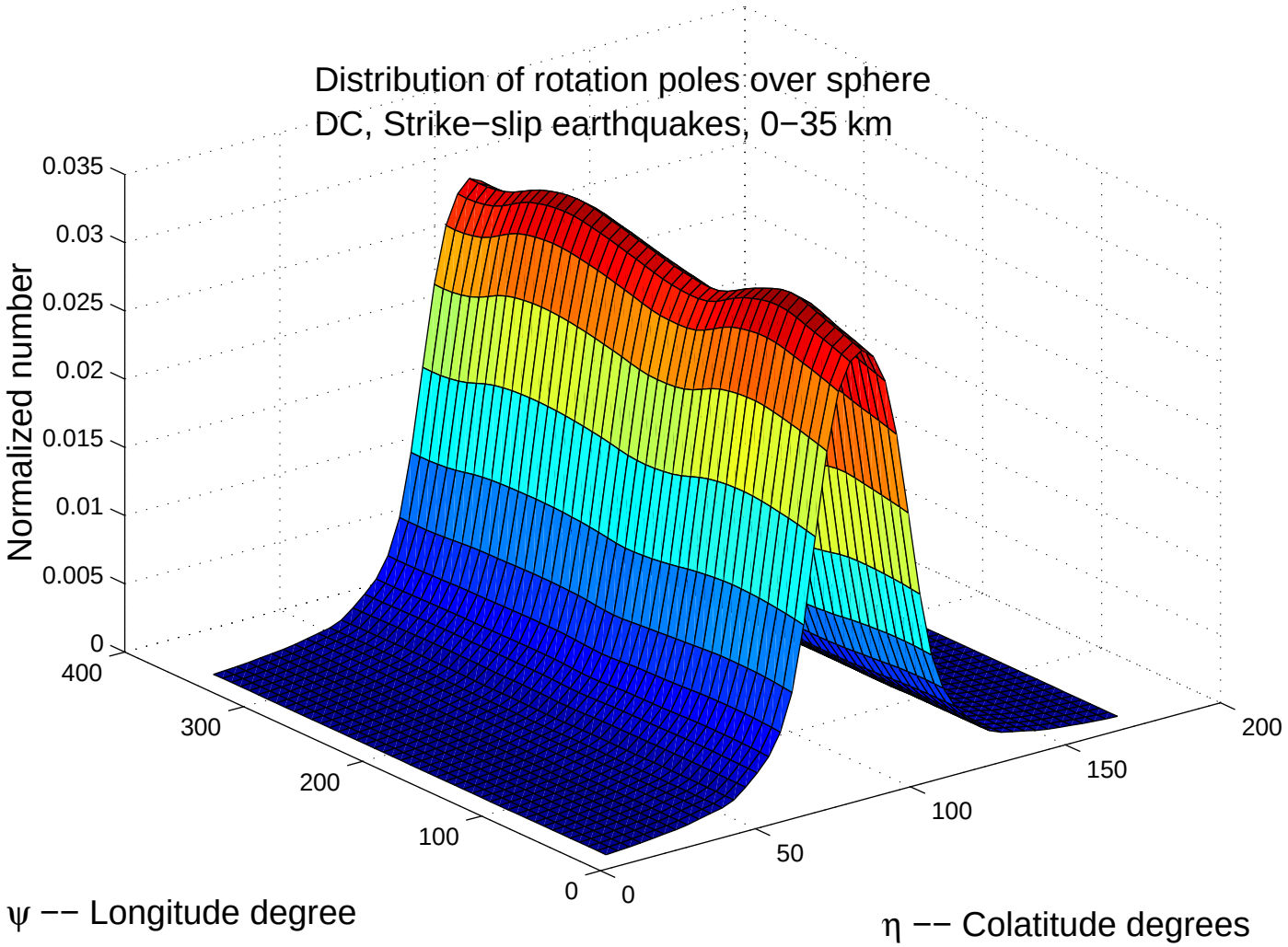
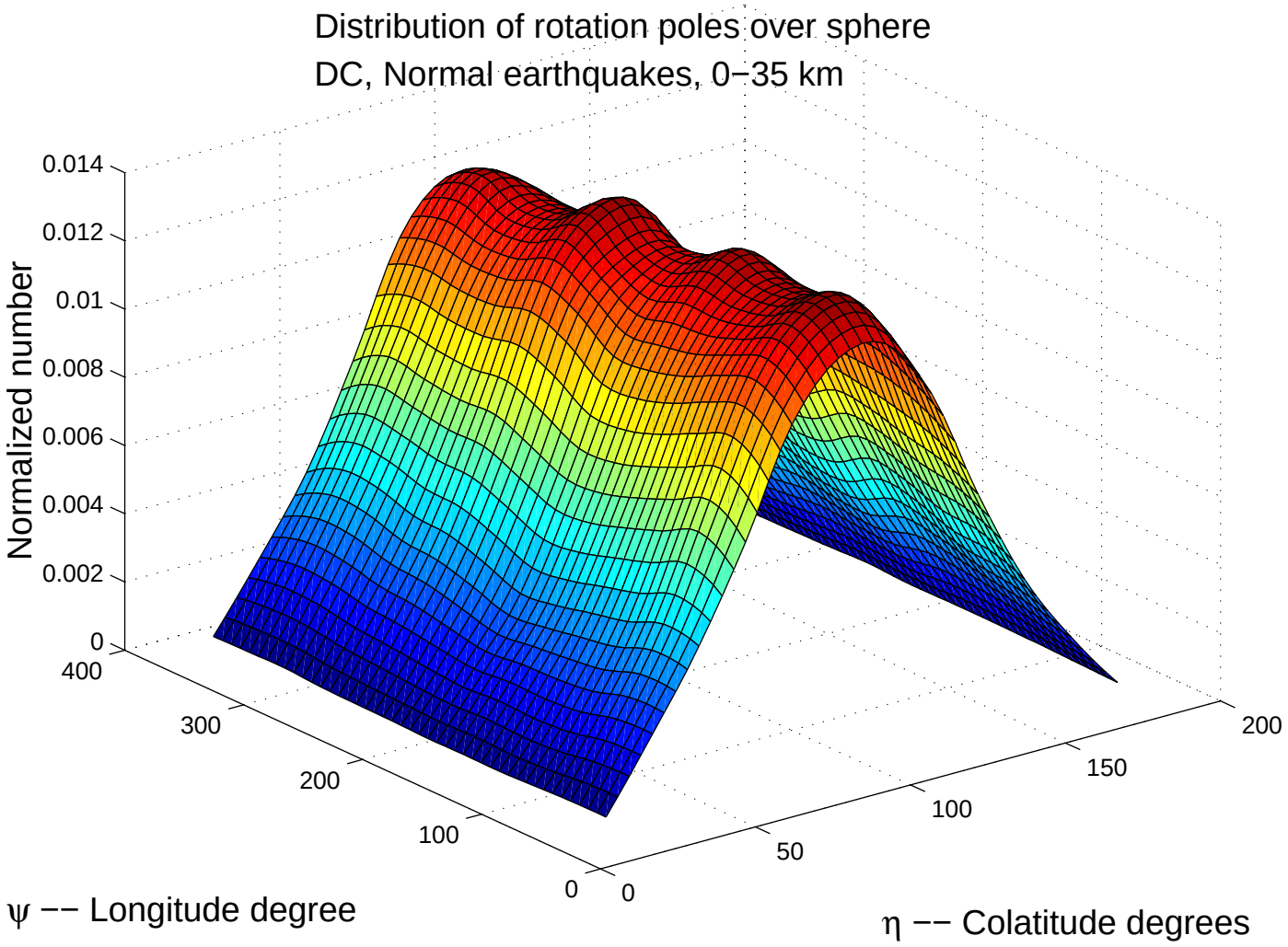


Fig. 21b: CMT−−GS−MT Rotation distr. hist., CMT−strike−slip, m>=5.3, 0−35 km




Distribution of rotation poles over sphere
DC, Thrust earthquakes, 0−35 km
Normalized number
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
400
300
200
100
0
0
50
100
150
200
ψ −− Longitude degree
η −− Colatitude degrees


Distribution of rotation poles over sphere
DC, Strike−slip earthquakes, 0−35 km
Normalized number
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
400
300
200
100
0
0
50
100
150
200
ψ −− Longitude degree
η −− Colatitude degrees


Distribution of rotation poles over sphere
DC, Normal earthquakes, 0−35 km
Normalized number
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
400
300
200
100
0
0
50
100
150
200
ψ −− Longitude degree
η −− Colatitude degrees