Chapter 2 Characterizing ground motion envelopes as functions of magnitude, distance, and site

The next two chapters describe efforts to predict and characterize envelopes of various channels of ground motion (vertical and horizontal acceleration, velocity, and filtered displacement). In this chapter, I describe the database, introduce a new parameterization of observed ground motion envelopes, and develop envelope attenuation relationships that predict ground motion envelopes as a function of time, given magnitude, distance, and site classification. In the following chapter, I use these envelope attenuation relationships to study some average properties of ground motion.

2.1 Database

The database consists of about 30,000 records of vertical and horizontal (both East-West and North-South) acceleration, velocity, and filtered displacement recorded at sites within 200 km of the source region of 70 Southern California earthquakes in the magnitude range $M2.0$ through $M7.3$. (The 1989 $M_w = 7.0$ Loma Prieta event is an exception, since it occurred in Northern California.) Included are data from the following large earthquakes: $M_w = 7.0$ Loma Prieta, 1991 $M_L = 5.8$ Sierra Madre, 1992 $M_w = 7.3$ Landers, 1992 $M_w = 6.4$ Big Bear, 1994 $M_w = 6.7$ Northridge main shock and $M_w = 5.1$ aftershock, and 1999 $M_w = 7.1$ Hector Mine. Events from the 2001 Anza sequence are also included.

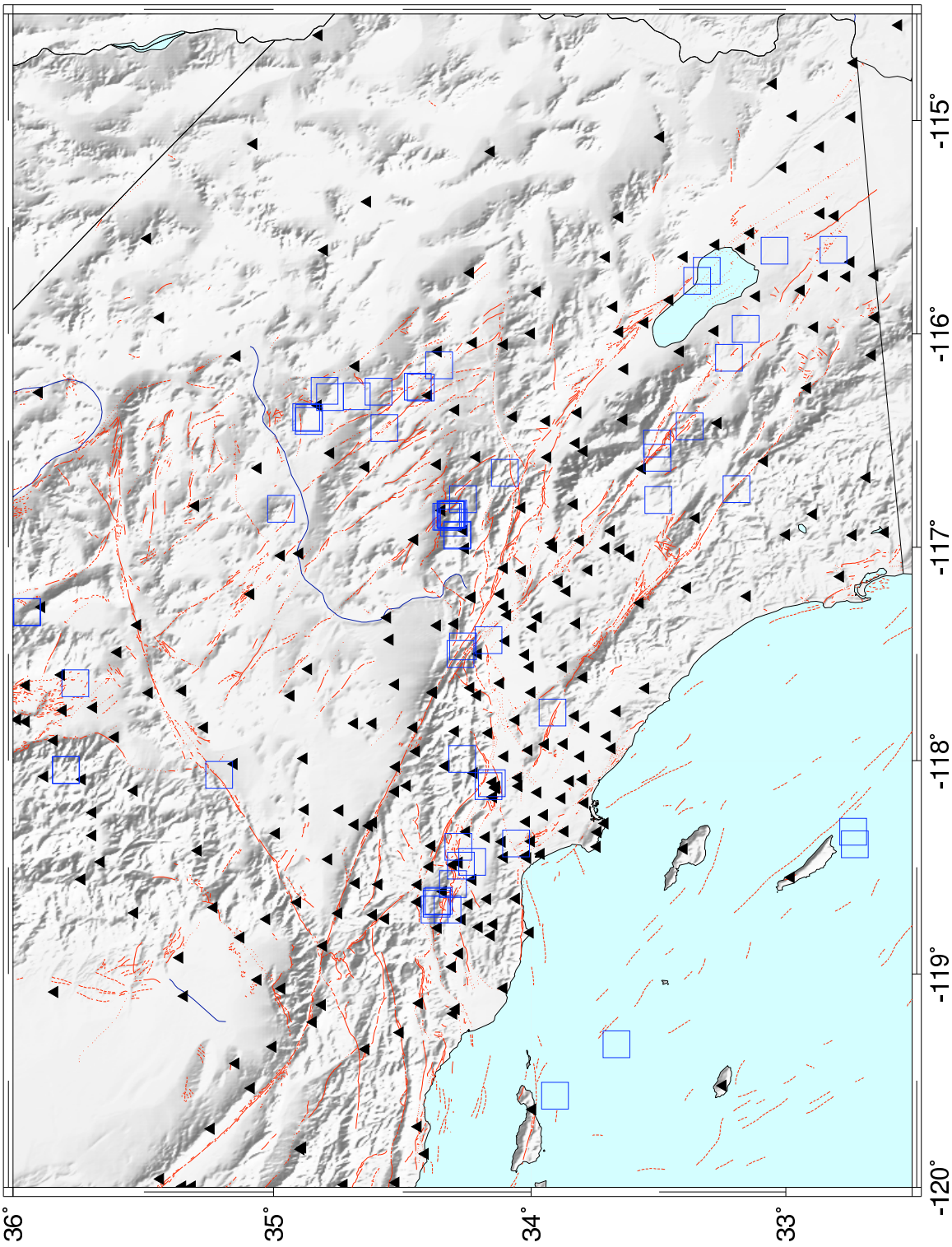

Figure 2.1: Map of (most) Southern California Seismic Network (SCSN) stations and earthquakes used in this study. Triangles denote stations, and open squares denote earthquakes. There are a few stations and earthquakes used in this study that are beyond the latitude limits of this map.

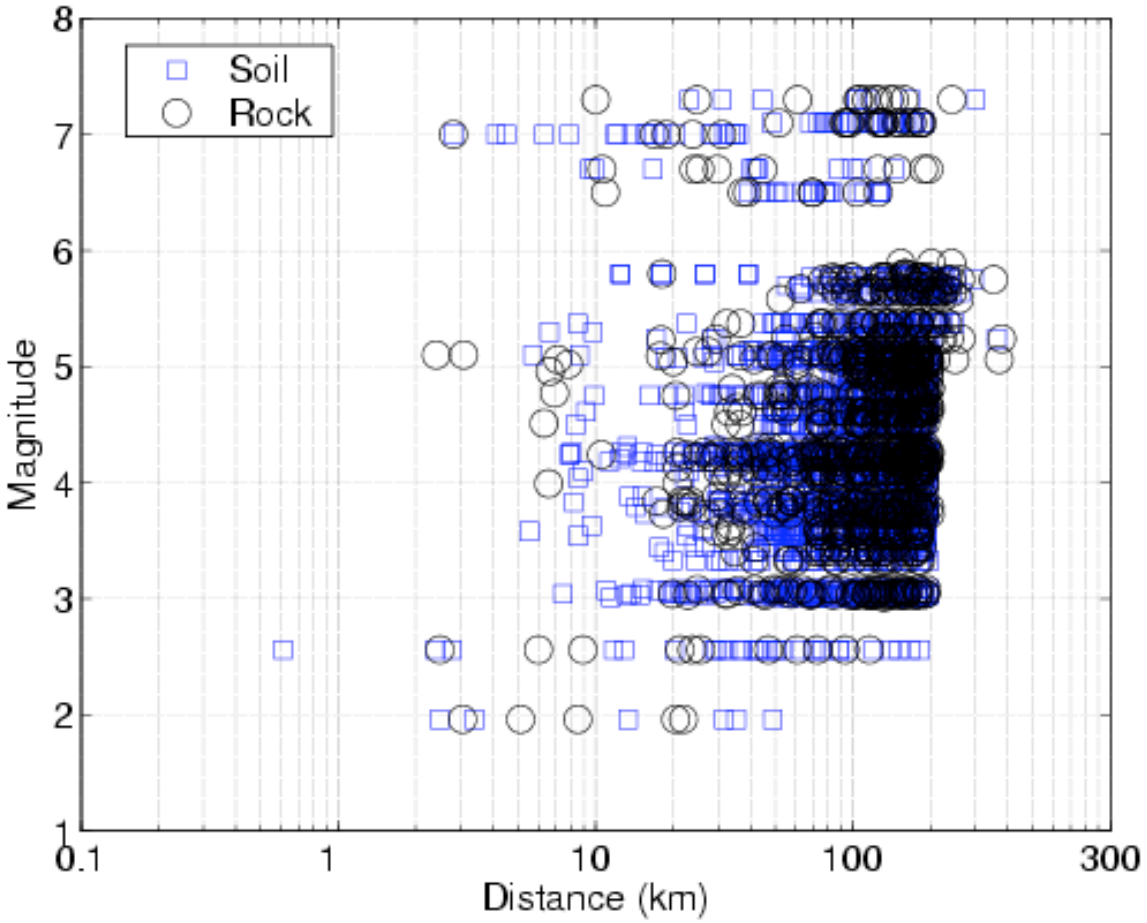

Figure 2.2: Distribution of records in database with magnitude and distance for a given channel of ground motion. This plot distinguishes between records obtained from rock sites (circles) and soil sites (squares). The distinction between rock and soil sites will be discussed in further detail later in this chapter.

2.1.1 Data processing

Whenever possible, I used data from digital Southern California Seismic Network (SCSN) stations. Typically, I downloaded data from the 100 sample per second, high gain, broadband (HH) channel, applied a baseline correction, and corrected for the instrument gain to get ground motion velocity. I differentiated the velocity records once to obtain ground motion acceleration, and integrated once to obtain ground motion displacement. When the HH channel was clipped, I downloaded the 100 sample per second, low gain accelerometer channel, applied a baseline correction, and accounted for the instrument gain to obtain ground motion acceleration. I integrated the ground motion acceleration once to get ground motion velocity, and integrated twice to get ground motion displacement.

I supplemented the SCSN records with data from strong motion networks operated by the United States Geological Survey (USGS), the California Geological Survey (CGS), and the California Strong Motion Instrumentation Program (CSMIP). The COSMOS Virtual Data Center (*http://db.cosmos-eq.org*) facilitated access to such strong motion data. Some of the earthquakes with COSMOS data include: 1989 $M_w = 7.0$ Loma Prieta, 1991 $M_L = 5.8$ Sierra Madre, 1992 $M_w = 7.3$ Landers, 1992 $M_w = 6.5$ Big Bear, 1994 $M_w = 6.7$ Northridge main shock and $M_w = 5.1$ aftershock, and 1999 $M_w = 7.1$ Hector Mine events.

Filtering displacement

Displacement records obtained via the processing described above can have significant low-frequency components. Errors in baseline corrections on acceleration and velocity introduce low-frequency errors in displacements obtained via integration. In larger earthquakes, the source of low-frequency energy is related to the finiteness of the rupture area (as opposed to the point source behavior of smaller earthquakes). The increase of low-frequency energy with magnitude is in fact the basis of several methods for seismic early warning (Allen and Kanamori, 2003; Nakamura, 1988), including a new method for quick magnitude estimation using ratios of ground motion

which is discussed in Chapter 4 of this thesis. Static offsets from large earthquakes also contribute a DC component to ground motion displacement. Smaller earthquakes involve smaller patches of slip, and radiate higher frequency energy than larger earthquakes (Allen and Kanamori, 2003). In smaller earthquakes, the low frequency energy in the displacement records is usually not related to the source process of the earthquake being examined, but rather from microseisms, which have energy peaked in the $5 - 9$ second period range. Microseisms can dominate the displacement record from smaller earthquakes, and filtering is necessary to obtain the signal caused by the earthquake being studied.

For seismic early warning purposes, we ideally would not filter out the lower frequency energy in the displacement record, since this frequency range contains information useful in discriminating between small and large earthquakes. However, without some high-pass filtering, microseisms dictate a lower bound on the magnitude range we can examine; below about $M3.5$, microseisms typically dominate the unfiltered displacement record.

To extend the magnitude range for this study, I applied a 3-second, 4-pole high-pass Butterworth filter to all displacement time histories in the database. This removes the constraint on the lower magnitude bound due to the microseisms; the smallest event in included in this study is $M1.95$. Fortunately, there is enough long period energy from larger earthquakes in the passband to distinguish between small and large events. The magnitude estimation method using ratios of ground motions to be described in Chapter 4 uses high-pass filtered displacements.

2.2 Envelopes of ground motion

After processing the raw data to get acceleration, velocity, and displacement time histories, I obtain ground motion envelopes by taking the maximum absolute amplitude of the ground motion time history over a 1-second window. It is this maximum over 1-second data stream that is available in closest to real-time from digital SCSN stations, making it an logical choice for seismic early warning applications. A signifi-

cant portion of this thesis is dedicated to characterizing and predicting this particular data stream as a function of magnitude, distance, and site condition.

Figure 2.3 shows the acceleration time history recorded at SCSN (then TriNet) station DGR during the 1994 Northridge mainshock. Figure 2.4 shows its associated envelope.

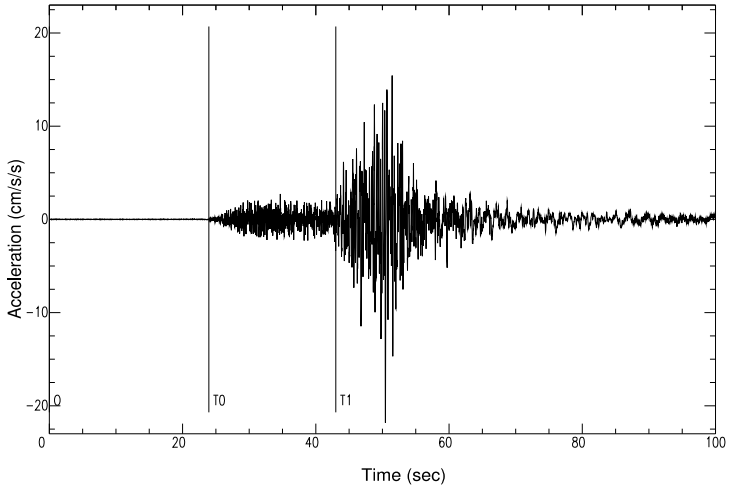

Figure 2.3: Acceleration time history recorded at station DGR during the 1994 $M_w = 6.7$ Northridge mainshock. Closest distance to the fault is 149.5 km, from (http://db.cosmos eq.org,). The P- and S-wave arrival times are marked by "T0" and "T1".

2.2.1 Parameterization of ground motion envelopes

I model the observed ground motion envelope as a combination of P-wave, S-wave, and ambient noise envelopes. The P-wave, S-wave, and ambient noise envelopes of a given record combine according to the rule:

$$E_{observed}(t) = \sqrt{E_P^2(t) + E_S^2(t) + E_{ambient}^2} + \epsilon \tag{2.1}$$

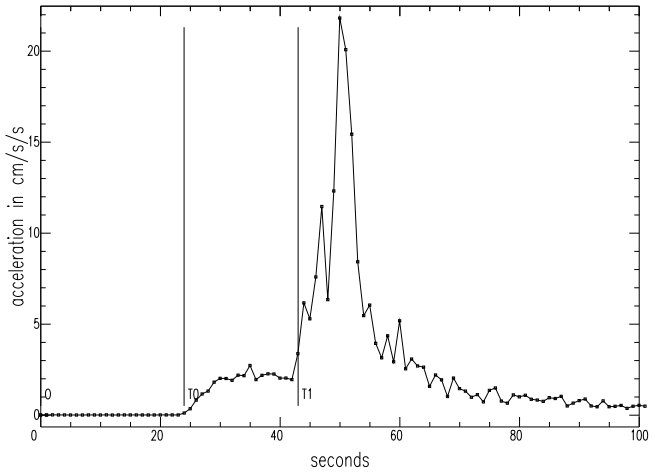

Figure 2.4: Envelope of acceleration time history recorded at station DGR during the Northridge mainshock.

where

$$
\begin{aligned}
E_{observed}(t) &= \text{envelope of observed ground motion} \\
E_P(t) &= \text{envelope of P-wave} \\
E_S(t) &= \text{envelope of S-wave and later-arriving phases} \\
E_{ambient} &= \text{ambient noise at the site} \\
\epsilon &= \text{difference between predicted and observed envelope}
\end{aligned}
$$

There are a number of ways to justify Eq. 2.1, and these are discussed in Appendix A.

The ambient noise, $E_{ambient}$, for a given time history is modeled as a constant. The P- and S-wave envelopes, $E_P(t)$ and $E_S(t)$, are each described by a rise time (t_{rise_P}, t_{rise_S}), a constant amplitude (A_P, A_S), a duration $(\Delta t_P, \Delta t_S)$, and two decay parameters (γ_P, γ_S), (τ_P, τ_S). Note: I found that a single decay parameter would typically fit the overall coda, but always had large misfits immediately after the peak P- or S-wave amplitudes. Jennings et al. (1968) also use two decay parameters

to describe the decay of envelope amplitudes with time following the peak ground motion. Using two decay parameters improves the fit between the modeled and observed envelopes at the cost of introducing trade-offs in the parameterization.

$$E_{i,j}(t) = \begin{cases} 0 & , \quad t < T_i \\ \frac{A_{i,j}}{t_{rise_{i,j}}}(t - T_i) & , \quad T_i \leq t < T_i + t_{rise_{i,j}} \\ A_{i,j} & , \quad T_i + t_{rise_{i,j}} \leq t < T_i + t_{rise_{i,j}} + \Delta t_{i,j} \\ A_{i,j} \frac{1}{(t - T_i - t_{rise_{i,j}} - \Delta t_{i,j} + \tau_{i,j})^{\gamma_{i,j}}} & , \quad t \geq T_i + t_{rise_{i,j}} + \Delta t_{i,j} \end{cases} \tag{2.2}$$

where

$$\begin{aligned} i &= \text{P-, S-wave} \\ T_i &= \text{P-, S-wave arrival times} \\ j &= \begin{matrix} \ddot{u}_{vertical} & , & \ddot{u}_{E-W} & , & \ddot{u}_{N-S} \\ \dot{u}_{vertical} & , & \dot{u}_{E-W} & , & \dot{u}_{N-S} \\ u_{vertical} & , & u_{E-W} & , & u_{N-S} \end{matrix} \end{aligned}$$

with $\ddot{u}$ denoting acceleration

$\dot{u}$ denoting velocity

u denoting displacement

A total of 11 envelope parameters (5 each for the P- and S-wave envelopes, and 1 constant for the ambient noise) are used to describe an observed ground motion envelope. (See Figure 2.5.)

2.2.2 Nonlinearity of parameterization

The main benefit of the above parameterization is that it permits a separate characterization for P- and S-waves. It makes intuitive sense that each of the body wave envelopes have a rise time, an amplitude with a finite duration, and parameters de-

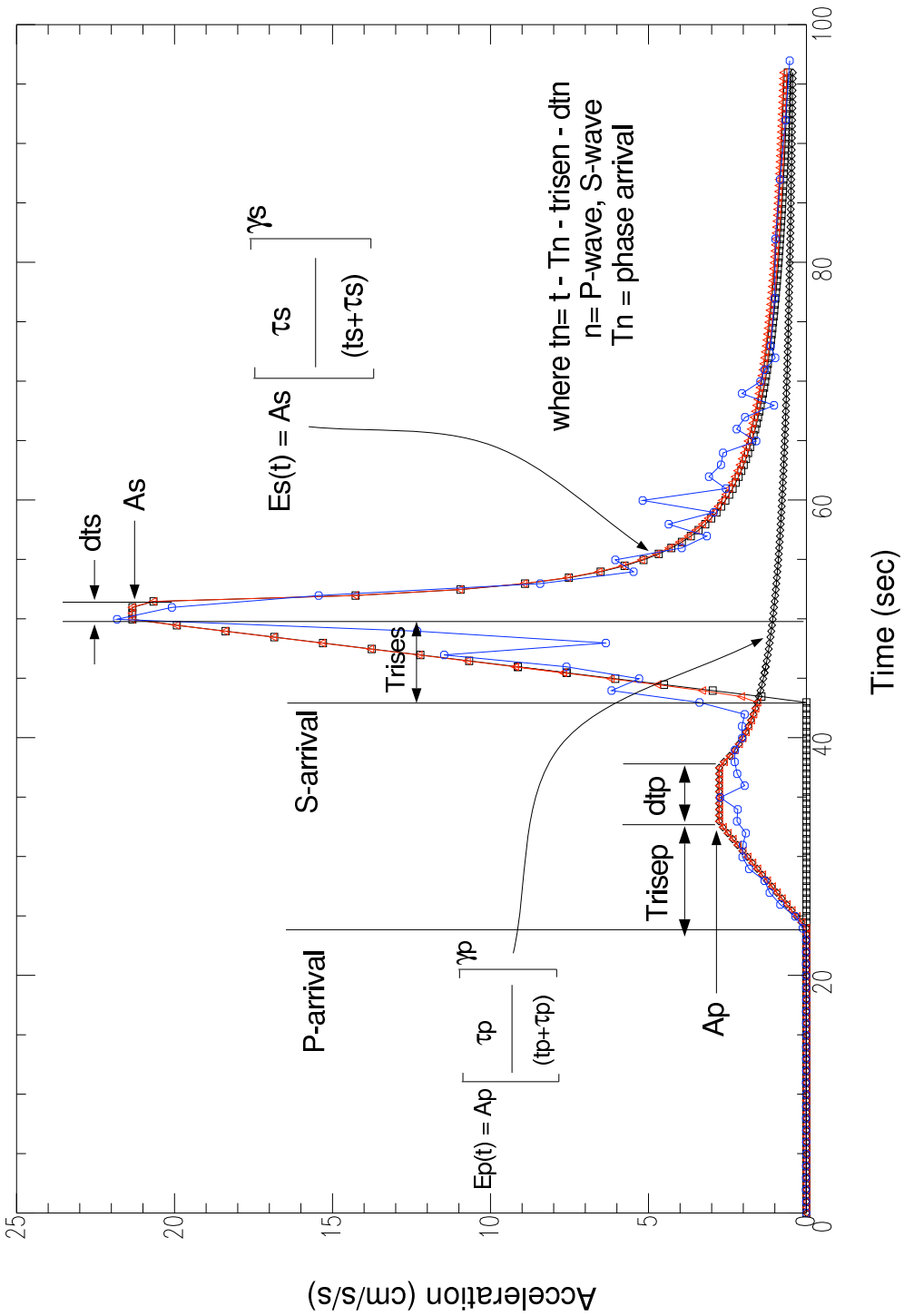

Figure 2.5: The observed ground motion envelope is decomposed into P-wave, S-wave, and ambient noise envelopes. Each of the P- and S-wave envelopes is described by 5 parameters (rise time, duration, amplitude, and 2 decay parameters). The Neighbourhood Algorithm is used to find the set of 11 envelope parameters that best fit the observed envelope in a least squares sense.

scribing its coda. Unfortunately, this "intuitive" parameterization comes at a cost. The piecewise linearity with time, and the Omori-type decay of the P- and S-wave coda, make finding the "best" 11 envelope parameters (in a least squares sense) for a given observed envelope a nonlinear problem. There are trade-offs between various parameters, for example, between the rise time and the duration and between the two decay parameters. Additional difficulties arise with the P-wave parameters at close distances. At distances less than about 20 km, there is less than 3 seconds of P-wave data before the S-wave arrival, making it difficult to constrain or estimate the P-wave decay parameters.

The aim is to quantify how envelopes of ground motion (such as Figure 2.3) depend on magnitude, distance, and site condition. In principle, we could postulate how the various envelope parameters depend on magnitude, distance, and site, and, along with Eqn. 2.2, find the model parameters that best fit all envelope time histories in our database in one single (very large) inversion. This is illustrated in Figure 2.6. However, this inversion would be highly nonlinear and quite large.

Instead, we pursue an iterative approach (Figure 2.7) wherein the single, very large nonlinear problem is replaced with very many small nonlinear inverse problems. In this iterative approach, the Neighbourhood Algorithm, a direct search method for nonlinear inversion developed by Sambridge (1999a, 1999b), is used to find the "best" 11 envelope parameters to fit each observed envelope history in the database. (Appendix B contains a brief description of the Neighbourhood Algorithm.) The parameterization described by Eqns. 2.1 and 2.2 is applied to all ground motion envelopes in the database. Envelope attenuation relationships that describe the magnitude, distance, and site dependence of the various envelope parameters are then developed. These envelope attenuation relationships are used to obtain more informed starting models for a second iteration of fitting all ground motion envelopes in the database and developing attenuation relationships for the envelope parameters. Figure 2.5 shows an observed envelope and the "best" P-wave, S-wave, and ambient noise decomposition using the Neighbourhood Algorithm.

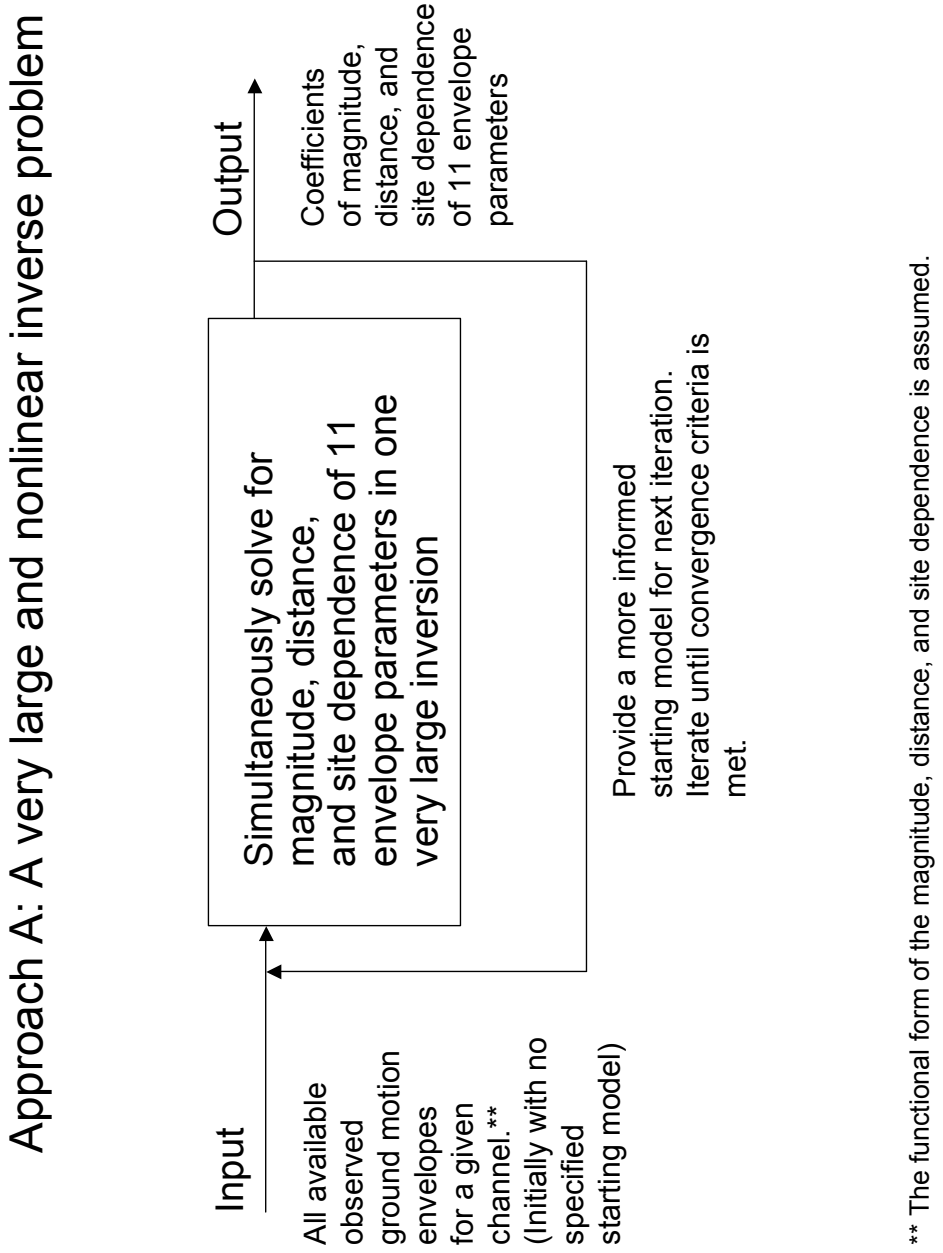

Figure 2.6: The aim is to obtain the magnitude, distance, and site dependence of ground motions envelopes. Given the database of observed envelopes and postulated forms for the magnitude, distance, and site dependence of various envelope parameters, this can be done in a single, very large, and very nonlinear inverse problem.

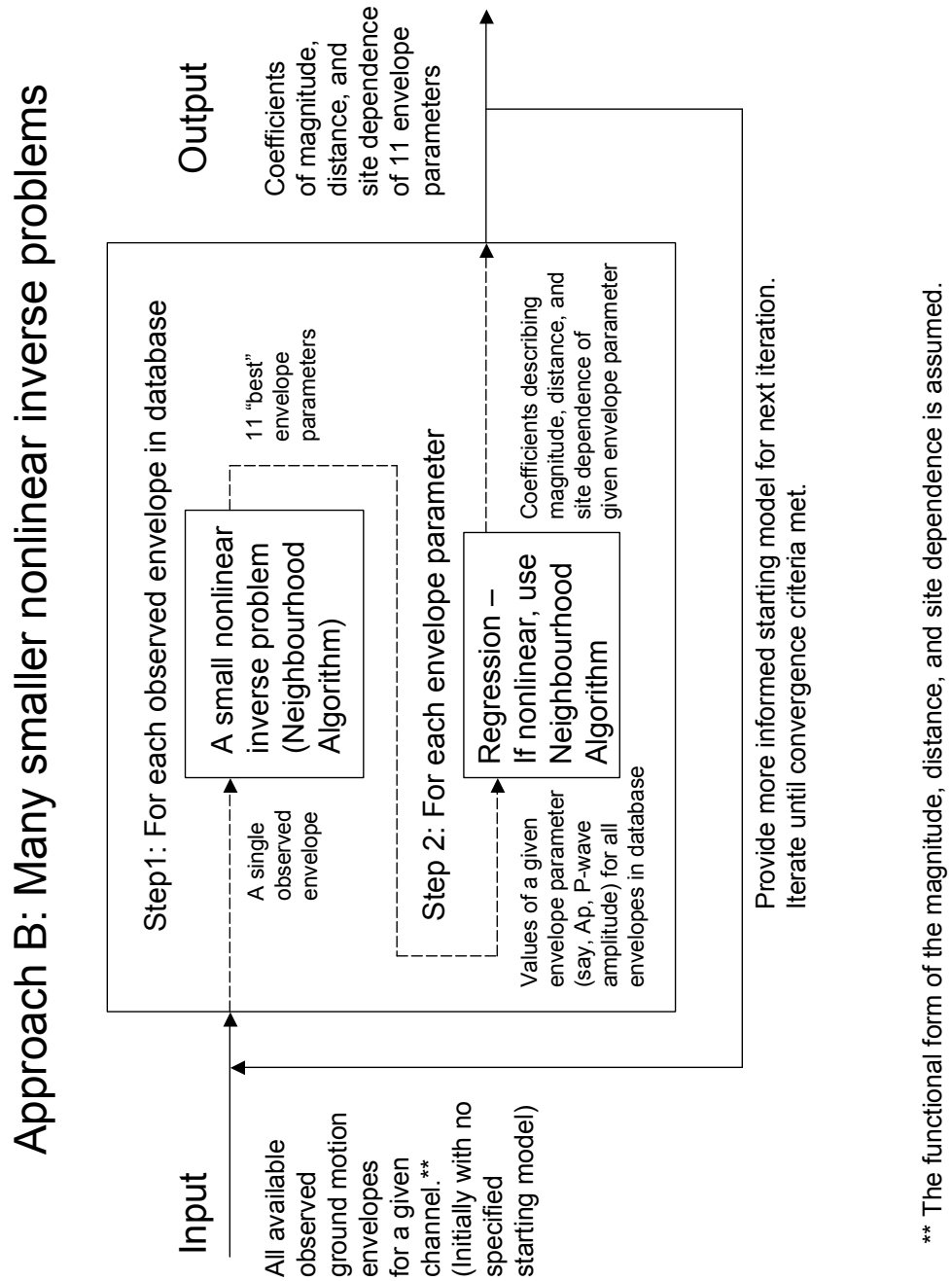

Figure 2.7: An iterative approach is used to obtain the magnitude, distance, and site dependence of various envelope parameters. The single, very large nonlinear inversion shown in Figure 2.6 is replaced by very many, relatively smaller nonlinear inversions.

2.3 Envelope attenuation relationships

One of the objectives of this research is to develop relationships that predict the envelopes of various channels of ground motion given magnitude, distance, and information about site condition. In this thesis, these functions are referred to as "envelope attenuation relationships", to distinguish them from more traditional strong motion attenuation relationships. I applied the envelope parameterization discussed to acceleration, velocity, and filtered displacement records in the vertical, East-West, and North-South directions. I developed envelope attenuation relationships for peak vertical and the root mean square of the peak horizontal acceleration, velocity, and displacement amplitudes. For each of these 6 channels (vertical and rms horizontal acceleration, velocity, and displacement), separate envelope attenuation relationships are developed for P- and S-wave amplitudes, and for rock and soil sites. Appendix C contains envelope attenuation relationships for various channels of ground motion. The envelope attenuation relationship for a given channel describes the magnitude, distance, and site dependence of each of the 11 envelope parameters (rise time, amplitude, duration, and 2 decay parameters for each of P- and S-wave envelopes, and constant for ambient noise). In this section, the development of these envelope attenuation relationships for the root mean square of the acceleration amplitudes recorded on rock sites of the two horizontal channels is described in a fair amount of detail. The development process is identical for the other channels of ground motion. Before beginning this discussion, a few words about existing strong motion relationships are in order.

2.3.1 A brief description of traditional strong motion attenuation relationships

Strong motion attenuation relationships have long been an important tool in both deterministic and probabilistic seismic hazard analysis. In estimating and designing engineered structures for the earthquake hazard at a given site, it is necessary to have relationships describing the expected ground motions at a given site as a func-

tion of factors such as source magnitude, distance away from the source, and local site characterization. Almost all attenuation relationships describe expected peak ground motions as a functions of earthquake magnitude, a measure of distance to the source region, and local site characterization. In standard regression terminology, the peak ground motion quantities are the dependent variables, and magnitude, distance, and site condition are the independent or predictor variables. Some attenuation relationships include other factors such as style of faulting (Boore, Joyner, and Fumal, 1997) and directivity effects (Somerville et al, 1997) to model the observed data. Typically, the quantities being modeled are peak acceleration, peak velocity, peak displacement, and response spectral quantities. Since it has been observed that there are systematic differences in ground motion characteristics across different tectonic regimes, attenuation relationships are typically characterized as being valid for one of three categories: shallow crustal earthquakes in active tectonic regions, shallow crustal earthquakes in stable continental regions, and subduction zone earthquakes (Abrahamson and Shedlock, 1997). Datasets for such attenuation relationships usually consist of global earthquakes of magnitude 5 and greater. Among the more widely used attenuation relationships are those developed by Boore, Joyner, and Fumal (1993, 1994, 1997), Abrahamson and Silva (1997), Campbell (1997), and Sadigh et al (1997). The Seismological Research Letters special issue on attenuation relationships (Volume 68, Number 1, Jan/Feb 1997) and a review paper by Campbell (2002) are good sources for more details about more traditional strong motion attenuation relationships.

2.3.2 Envelope attenuation relationships: amplitude-type parameters

The terms "ground motion model" and "attenuation relationship" are interchangeable. They refer to the same concept: a mathematical expression that relates ground motion quantities to other parameters describing the earthquake source, the wave propagation path between the source and the site, and the local conditions at the site (Campbell, 2002). Recall that the envelope parameterization (Eqn. 2.2) involves

a P-wave and and S-wave amplitude for each envelope history in the database. The functional form used to describe the magnitude, distance, and site dependence of the P and S wave amplitudes is given below:

$$
\begin{aligned}
\log_{10} A_{ij} &= a_i M - b_i(R_1 + C_i(M)) - d_i \log_{10}(R_1 + C_i(M)) + e_{ij} + \epsilon_i \qquad (2.3)\\
i &= 1, \ldots, 24 \text{ (P- and S-wave amplitudes on rock and soil sites for 6 channels)}\\
j &= 1, \ldots, \text{ number of stations}\\
A_{ij} &= \text{“best” ground motion envelope amplitude from Neighborhood Algorithm}\\
M &= \text{SCSN magnitude } (M_w \text{ for } M > 5.0)\\
R &= \text{ epicentral distance in km for } M < 5\\
&\quad \text{closest distance to fault for } M > 5.0 \text{ (when available)}\\
R_1 &= \sqrt{(R^2 + 9)}\\
C_i(M) &= (\arctan(M - 5) + 1.4)(c_{1i} \exp(c_{2i}(M - 5)))\\
e_{ij} &= \text{constant}_i + \text{station corrections}_{i,j}\\
\epsilon_i &= \text{statistical (or prediction) error, } \sim NID(0, \sigma^2)
\end{aligned}
$$

The A_{ij}s are the “best” amplitudes (P- or S-wave) from fitting Eqns. 2.1 and 2.2 to the observed ground motion envelopes in our database via the Neighborhood Algorithm. The $A_{ij}, M, R,$ and ϵ_{ij} are $(m \times 1)$ column vectors for the i^{th} channel, where m is the number of data points in a given regression. For each of the $i = 1 \ldots 24$ amplitude parameters, Eqn. 2.3 is nonlinear in the unknowns, which are $(a_i, b_i, c_{1i}, c_{2i}, d_i, e_{ij})$.

Eqn. 2.3 has strong influences from traditional strong motion attenuation relationships, in particular from the work of Boore and Joyner (1982), Boore, Joyner, and Fumal (1993, 1997), and Campbell (1981, 2002). The physical motivations for the various terms of Eqn. 2.3, paraphrasing both Boore and Joyner (1982) and Campbell (2002), are as follows:

- $\log_{10} Y \propto M$ is consistent with the basic definition of magnitude (Richter, 1935) as the logarithm of ground motion amplitude

- $\log_{10} Y \propto \log_{10}(R)$ is consistent with geometric attenuation of the seismic wave front away from the source

- $\log_{10} Y \propto R$ is consistent with anaelastic attenuation due to material damping and scattering as seismic waves propagate through the crust, which is manifested as an exponential decay of amplitudes with travel time

- $\log_{10} Y \propto e$, where e are site correction terms, is consistent with the multiplicative nature of site effects

The residuals or errors, ϵ are the difference between the observed values and those predicted by Eqn. 2.3.

$$\epsilon_k = \log_{10}(A_{obs,k}) - \log_{10}(A_{pred,k}(M,R)) \tag{2.4}$$

$$\sigma^2 = \frac{\sum_{k=1}^{n} \epsilon_k^2}{ndof} \tag{2.5}$$

where *obs* identifies the observations, and *pred* the predicted values; *ndof* is the number of degrees of freedom, n the total number of data points, and k is an index denoting the k^{th} data point in the database. It is assumed that the errors ϵ have zero mean with constant variance σ^2 and are independent and identically distributed. This is a standard assumption in regression analyses. For linear models, this assumption allows the use of statistical significance tests and the development confidence and prediction intervals (Rawlings et al, 1998). Since Eqn. 2.3 is nonlinear, we cannot perform formal significance tests. Nonetheless, the validity of this assumption has important implications in the seismic early warning applications discussed in Chapter 4.

Note: Bayes' theorem is used in this thesis in parameter estimation related to seismic early warning (Chapters 4 and up). However, it can also be used in model selection applications. Bayes' theorem in model selection (finding the most probable

model class given data and a set of possible model classes) has been described by Beck and Yuen (2004), and has applications in developing attenuation relationships.

Magnitude-dependent saturation of ground motion amplitudes

The $C(M)$ term in Eqn. 2.3 is adopted from Campbell's work on near-source attenuation of peak acceleration (1981). His original saturation term is

$$C(M) = c_1 \exp(c_2(M)) \tag{2.6}$$

This term allows ground motion amplitudes at close distances to the fault to be independent of magnitude for large ($M > 5$) events, and follows the recommended modification to attenuation curves made by Hadley and Helmberger (1980). Based on their simulations of strong ground motion, Hadley and Helmberger found a significant decrease in the slope of attenuation curves at smaller epicentral distances. That is, simulations indicate that the shape of the attenuation curves is magnitude-dependent, with ground motion amplitudes in the near-source region of large earthquakes approaching a limiting value. Campbell was the first to find empirical evidence for such saturation in his study of near-source attenuation of peak acceleration (1981). His dataset consisted of near-source (within 50 km) data from global earthquakes with magnitudes larger than 5. In contrast, the database used in this work contains records within 200 km of events in the magnitude range $M2.0$ through $M7.3$. The possible mechanisms for saturation effects are 1) nonlinear soil response when subject to ground motions above a certain threshold, and 2) geometric considerations - the distance decay due to a point source becomes distance independent as the point expands to a plane (Hadley and Helmberger, 1980). These saturation factors become important for $M > 5$ events, when 1) source dimensions become comparable to station distances, and 2) ground motion amplitudes become large enough to induce yielding in soils. I modify Campbell's saturation term (Eqn. 2.6) with an $(\arctan(M-5)+1.4)$ term

$$C(M) = (\arctan(M-5) + 1.4)(c_1 \exp(c_2(M-5))) \tag{2.7}$$

This "turns on" the saturation effects when $M > 5$, but allows the logarithm of ground motion amplitude to scale linearly with magnitude for $M < 5$.

After preliminary analyses, strong trade-offs were found between the saturation terms c_1 and c_2. To resolve these trade-offs, the range of c_2 is limited to be ~ 1, and c_1 is allowed to vary between 0 and 3. Values of c_1 close to 0 mean no saturation, with increasing values indicating stronger saturation effects. In Figure 2.8, C(M) is plotted as a function of magnitude for various values of c_1 (with the range of c_2 limited to be ~ 1).

Site classification

This study uses a binary rock-soil classification based on the NEHRP site classification scheme (7 classes, A through E), which is in turn based on the average shear-wave velocity in the upper 30 meters (V_S^{30}) of a site profile. For each of the stations used in this study, the station coordinates were used to associate with the station a particular site geology based on a generalized geologic map of Southern California produced by the California Geological Survey (Wills et al, 2000). The site geology is used to obtain an estimate of V_S^{30}, which is translated to a NEHRP site class, and then to a binary site class, as shown in Table 2.1. The binary scheme in this study denotes NEHRP site class BC and above as "rock", and NEHRP site class C and below as "soil". Table 2.1 lists the NEHRP classes, the mean V_{30} for the given geologic unit, and the associated binary classification. Separate attenuation relationships are derived for rock and soil classes. This allows us to address whether ground motions on rock and soil sites are significantly different over the magnitude and distance ranges covered by this study. Such comparisons are the subject of the following chapter. Since CISN stations contribute most of the ground motions in the dataset, this study does not include records from E class sites. (CISN stations are not typically located on E class sites.)

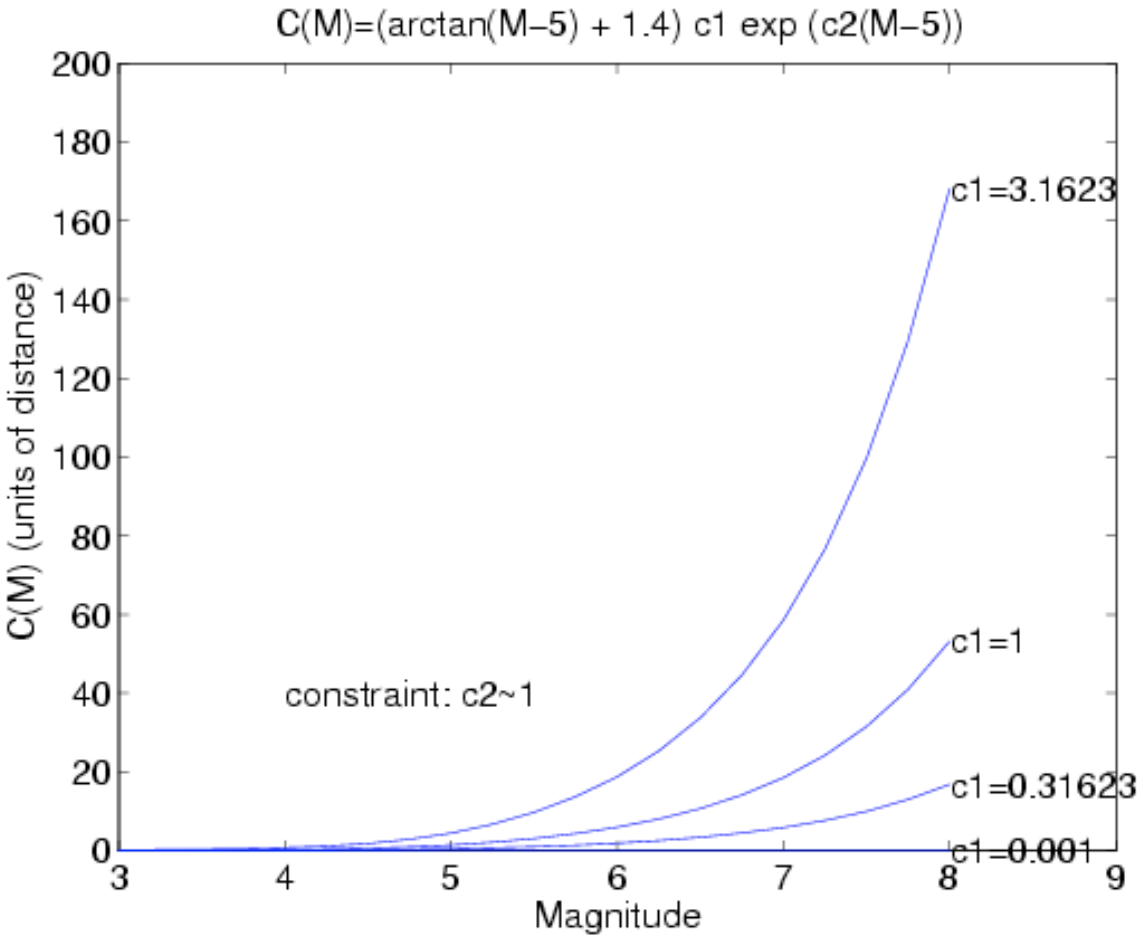

Figure 2.8: Magnitude-dependent saturation term for P- and S-wave envelope amplitudes. The range of c_2 is limited to be ~ 1, and c_1 is a parameter determined via regression. Values of c_1 close to 0 mean no saturation, with increasing values indicating stronger saturation effects. Note that in this formulation, $C(M)$ is in units of distance. That is, as $C(M)$ becomes large, it increases the "effective epicentral distance" of the station (see Eqn. 2.3).

NEHRP Class	mean V_S^{30} (m/s)	Binary Class
B	724	Rock
BC	686	
C	464	Soil
CD	372	
D	301	
E	163	

Table 2.1: NEHRP site classes, maximum V_S^{30} for given geologic unit, and binary class

Nonlinear regression

The functional form used to describe the magnitude, distance, and site dependence of the P- and S-wave envelope amplitude parameters is given by Eqn. 2.3. The Neighbourhood Algorithm (Sambridge, 1999a; Sambridge, 1999b) is used to find the model parameters $(a_i, b_i, c_{1i}, c_{2i}, d_i, e_i)$ that minimize the residual sum of squares between the weighted dataset and the estimates from Eqn. 2.3.

Figure 2.9 shows 1) the rms horizontal S-wave acceleration amplitude envelope parameter (without station corrections) from rock sites, and 2) curves of predicted S-wave acceleration versus distance for various prescribed magnitudes. Without saturation effects, the amplitude curves for the various magnitudes should have the same shape. The effect of the saturation term, $C(M)$, is to cause a flattening of the attenuation curves at close distances to large events. It appears that the maximum predicted S-wave acceleration is on the order of $0.5g$.

The standard error of regression is a measure of how well the model fits the data. The standard error of regression is given by

$$\sigma = \sqrt{\frac{\sum_{k=1}^{n} (A_{obs,k} - A_{pred,k})^2}{ndof}} \tag{2.8}$$

where n is the number of data points, k denotes the k^{th} data point in the regression, and $ndof$ is the number of degrees of freedom. $ndof$ is equal to the number of

data points less the number of model parameters to be determined in the regression. Without the station corrections, there are $(n-6)$ degrees of freedom, due to the 6 regression unknowns, $(a_i, b_i, c_{1i}, c_{2i}, d_i, e_i)$. The standard error of regression for rms horizontal S-wave acceleration amplitude without the station corrections is $\sigma = 0.31$.

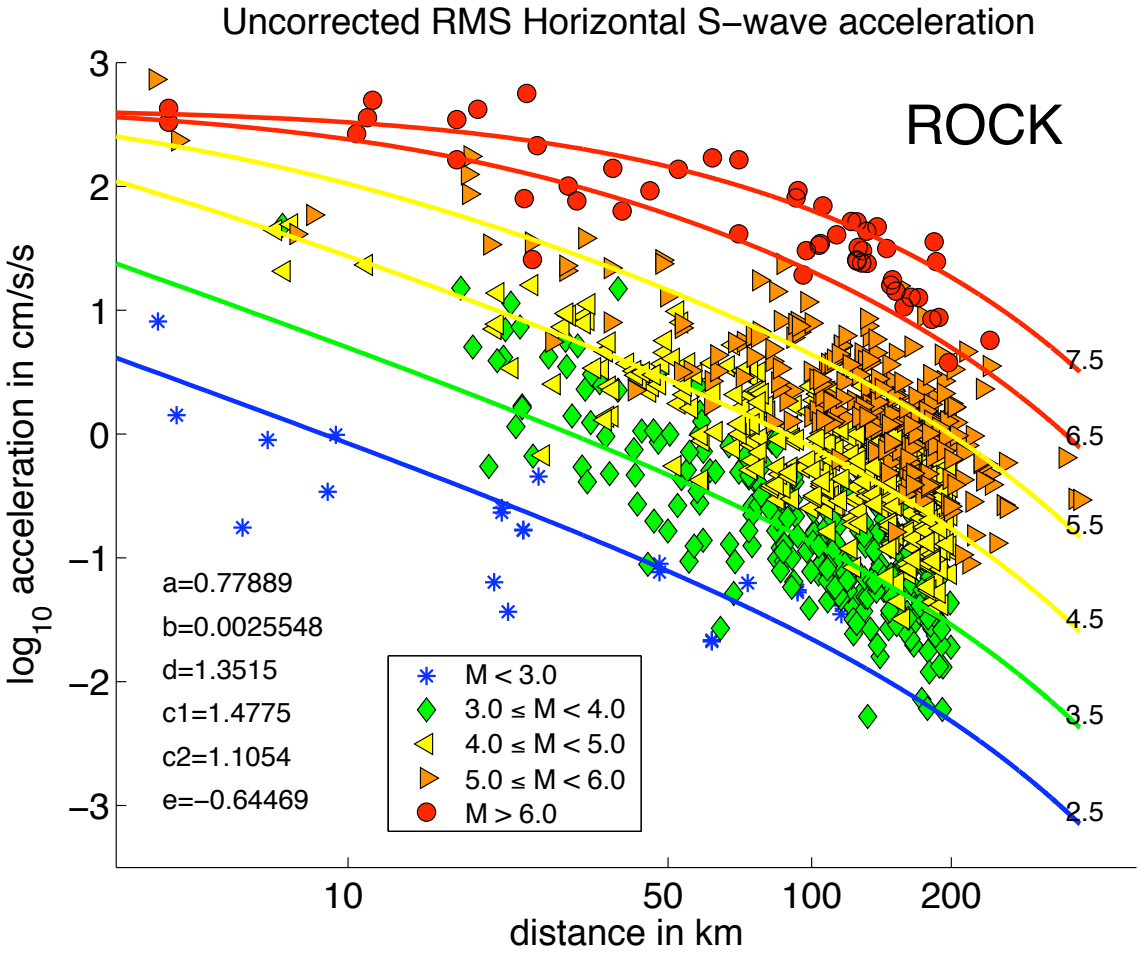

Figure 2.9: Uncorrected root mean square horizontal S-wave acceleration amplitude on rock sites. The standard error of regression is $\sigma = 0.31$. The symbols are the data. The solid lines are the amplitude levels predicted by the best fit attenuation relationship at the indicated magnitudes. The best fit model parameters are listed on the plot.

Station corrections

This study analyzed ground motions recorded by the seismic network in Southern California over a period of about 4 years. Some stations that were part of earlier predecessors of SCSN (TriNet and TERRAscope), such as Pasadena (PAS), Domenegoni Reservoir (DGR), and Victorville (VTV) have records from as many as 50 events. Due to the aftershock activity following the 1999 $M_w = 7.1$ Hector Mine earthquake, station Hector (HEC), also has a large amount of data. This gives us the opportunity to examine site amplification and to determine station-specific correction factors.

Station corrections are defined for all stations that recorded 5 or more events. Station corrections are the average residual at a given station relative to Eqn. 2.3. Rock stations have their corrections defined relative to the rock attenuation relationships; likewise for soil stations. Each channel (i.e., vertical acceleration, horizontal velocity, etc) has its own distinct set of station corrections.

If we want to account for station corrections in predicting ground motions, we can add the station correction (in $\log_{10}$ units) to the predicted ground motions. If we are analyzing data and want to account for station corrections, we subtract the station correction (in $\log_{10}$ units) from the observed amplitudes (in $\log_{10}$ units).

Station corrections for rms horizontal S-wave acceleration of the CISN rock stations used in this study are shown in Figure 2.10. This figure also shows the number of earthquakes recorded at the given station. It is useful to know how many records contribute to defining a correction term for a particular station; it is indicative of the statistical significance of the correction. For instance, stations PAS, PFO, and ISA have corrections in excess of -0.3 log units. This translates to a deamplification of more than 50% relative to the average rock station. The relatively large number of records at these stations (about 50, 20, and 10 records, respectively) indicate that the low station corrections are not due to randomness or chance (which might be suspected if the corrections were based on only one or two records), but are indicative of a consistent deamplification of ground motion at these sites. Incidentally, this approach allows us to identify "average" rock stations, which are defined as stations

at which observed ground motions are closest to those predicted by Eqn. 2.3. These are stations whose correction terms are close to zero. The "average" CISN rock stations include Goldstone (GSC), Palomar (PLM), Hector (HEC), Edwards (EDW), and AGA (Agoura Hills).

Figure 2.11 shows the rms horizontal station-corrected S-wave acceleration amplitude from rock sites and curves of predicted S-wave acceleration versus distance for various prescribed magnitudes. The station-corrected data points in Figure 2.11 cluster closer about the predicted attenuation curves (which are identical to the attenuation curves in Figure 2.9, since they are generated by the same model) than the uncorrected data (Figure 2.9). This indicates that we are better fitting the corrected data, which is expected, since station corrections introduce as many additional regression parameters as the number of stations. The total number of degrees of freedom ($ndof$ in Eqn. 2.8) in the uncorrected case is $(n-6)$, which is the number of unknown coefficients in Eqn. 2.3; $ndof$ is $(n-37)$ (n - 6 unknowns - 31 station corrections) when station corrections are included. The standard error of regression, with the station corrections is $\sigma_{corrected} = 0.24$, which is corresponds to $\sim 20\%$ reduction relative to the standard error in the uncorrected case ($\sigma_{uncorrected} = 0.31$).

Another way to evaluate the effect of station corrections is by plotting the observations against the predicted amplitudes with and without station corrections. The solid black line in Figure 2.12 is the true regression line, which has slope 1 and corresponds to the case of exactly predicting the observations. The crosses and circles denote the uncorrected and station-corrected cases. Station corrections improve predictions (by $\sim 20\%$); the circles are more closely clustered about the true regression line than the crosses.

Regression diagnostics

Eqn. 2.3 involves the assumption that the residuals (or errors) ϵ between the observed and predicted ground motion amplitudes are normally and independently distributed, with zero mean and constant variance σ^2. (Statistics text state this more succinctly as $\epsilon \sim NID(0, \sigma^2)$.) Here the validity of this assumption is demonstrated by plotting the

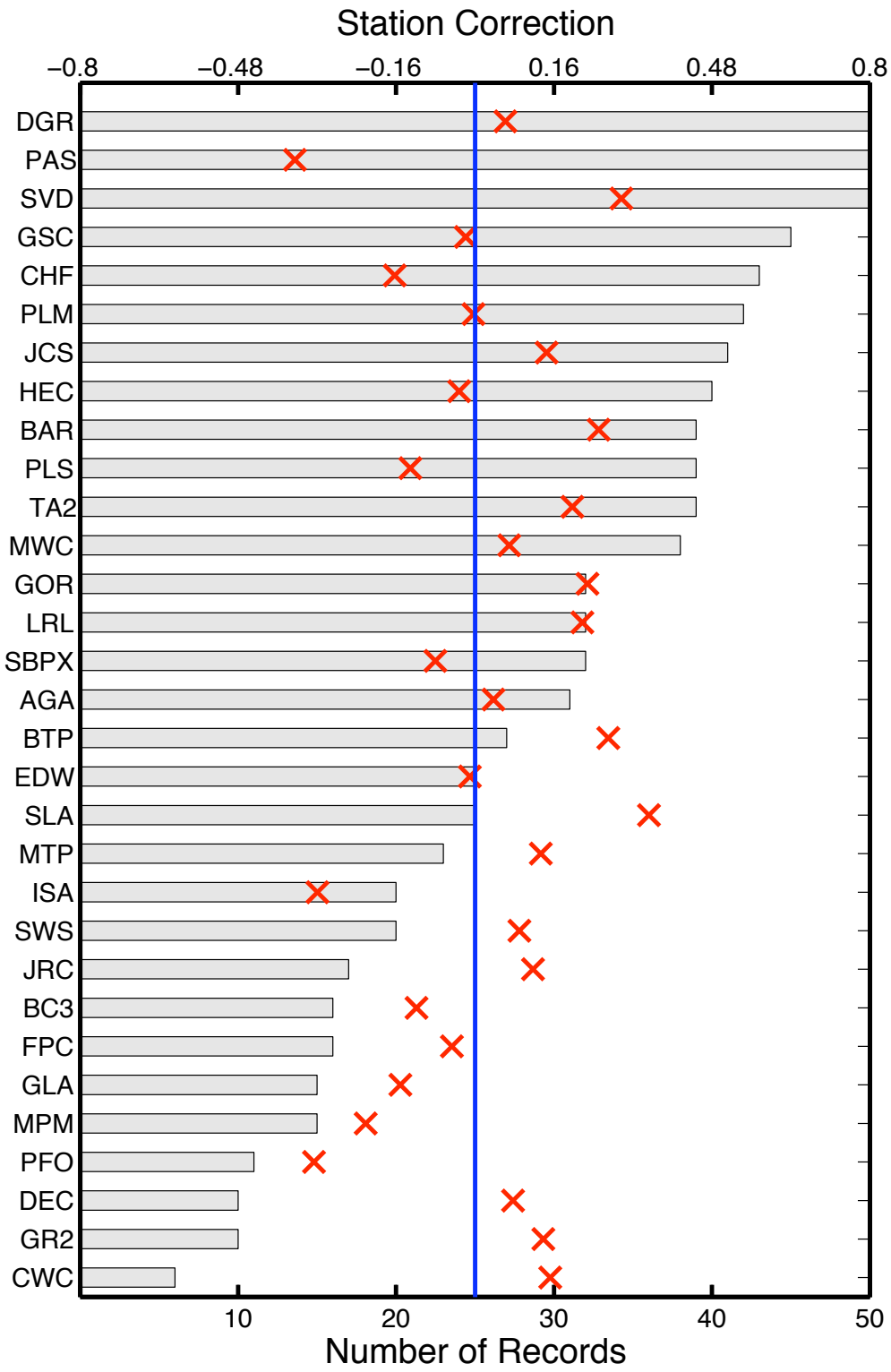

Figure 2.10: Station corrections for root mean square horizontal S-wave acceleration amplitude on rock. Corrections (denoted by "x", using the top x axis) are in $\log_{10}$ units. The bars indicate the number of records at a given station (bottom x axis).

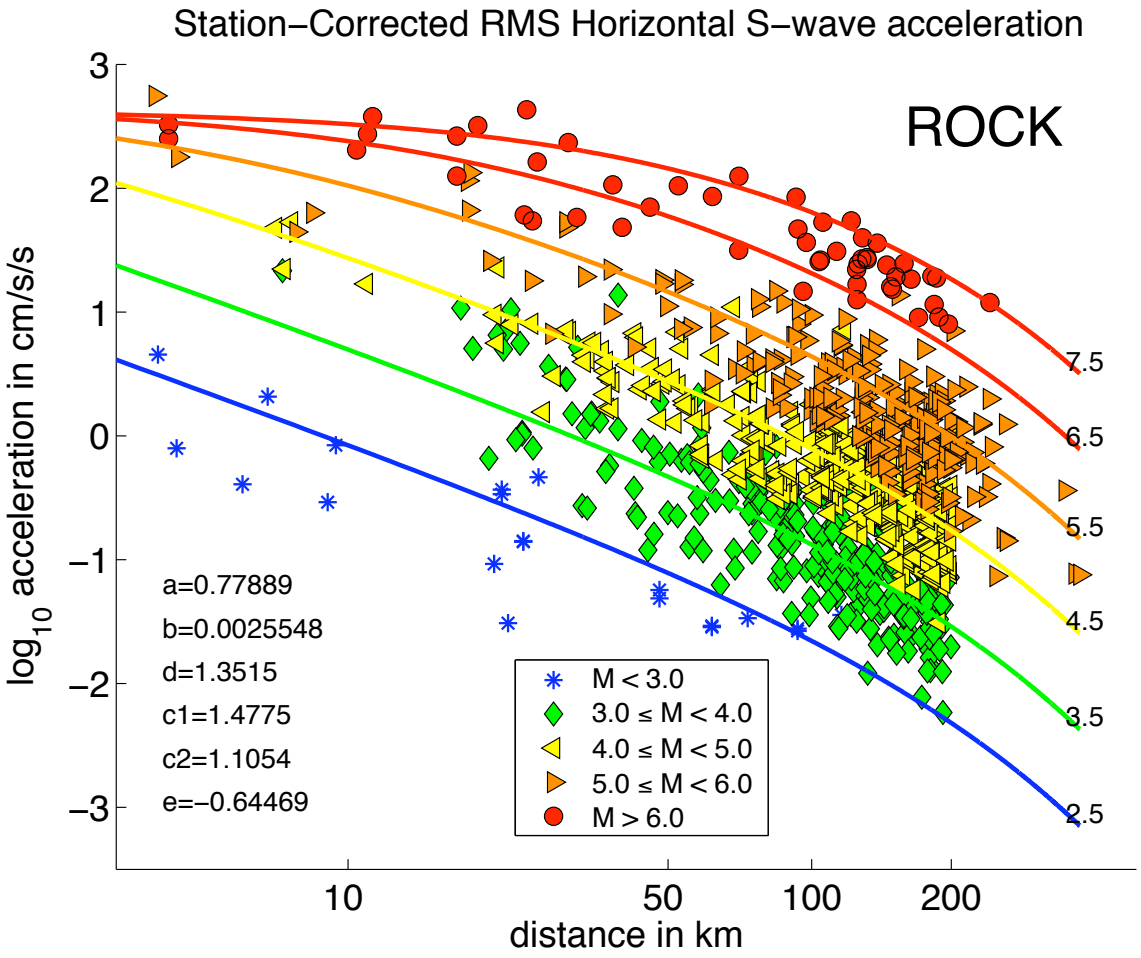

Figure 2.11: Station-corrected root mean square horizontal S-wave acceleration amplitude on rock. The standard error of regression is $\sigma = 0.24$, which is a has a reduction of $\sim 20\%$ relative to the standard error of regression without the station corrections.

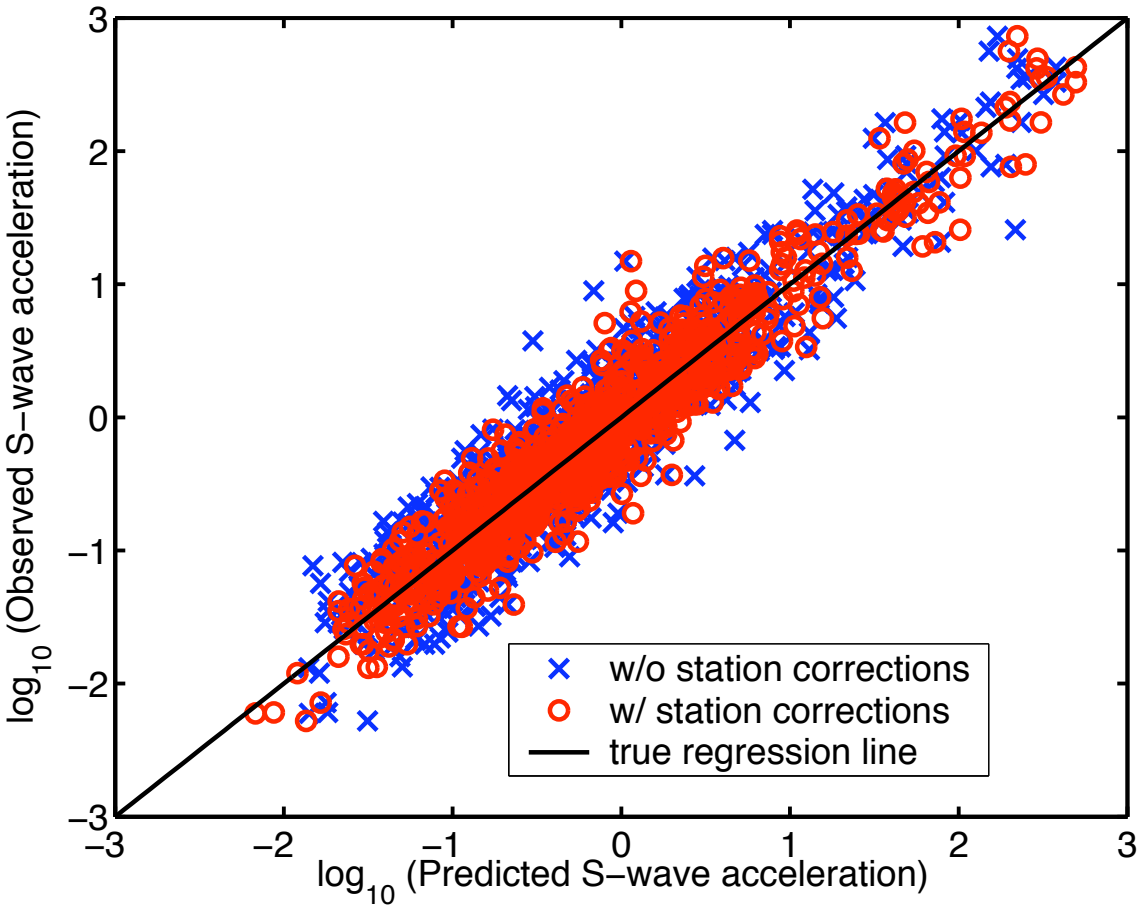

Figure 2.12: Comparison of predicted and observed S-wave acceleration amplitudes on rock. Observed amplitudes are plotted against predicted amplitudes. The solid black line, the true regression line, has slope 1 and corresponds to exactly predicting the observed data. Station corrections improve ground motion predictions; this is evident from the tighter clustering of the circles about the true regression line.

residuals against the predictor variables (magnitude and distance), and the predicted quantities (envelope amplitudes). The validity of this normality assumption becomes important in Chapter 4, when the amplitude attenuation relationships are used to quantitatively define the consistency of particular magnitude and distance estimates with the available observations in a seismic early warning application.

The errors or residuals ϵ should be simply noise, if we have adequately accounted for the magnitude and distance dependence of the ground motion amplitudes. Identifiable structure in the residuals is indicative of either a missing predictor variable or that a different distribution (for example, Poisson, binomial) is more appropriate for the data. If the assumption $\epsilon \sim NID(0, \sigma^2)$ is met, a plot of ϵ against the predictors (magnitude and distance) and the predicted values ($\log_{10}(A)$) should show a a random scattering of points about the $\epsilon = 0$ line, with most of the points within the 95% confidence intervals, $\epsilon \pm 2\sigma$) (Rawlings et al, 1998). Figures 2.13, 2.14, and 2.15 show residuals plotted against magnitude, distance, and predicted amplitude, respectively for the uncorrected (top) and station-corrected (bottom) cases. In these Figures, the solid lines correspond to $\epsilon = 0$; the dashed lines are the approximate 95% confidence intervals, given by $\pm 2\sigma$. Most points lie within these 95% confidence intervals. There appears to be no identifiable structure in the residuals with magnitude, distance, or predicted amplitude. Histograms of the residuals (Figure 2.16) show that they fit a normal distribution. In general, these Figures show that the normality assumption in Eqn. 2.3 is not being violated, and are indicative of the variance reduction due to station corrections.

2.4 Envelope attenuation relationships: rise time, duration, and decay parameters

Traditional strong motion attenuation relationships are well-documented in seismology and earthquake engineering literature. The envelope attenuation relationships for P- and S-wave amplitude parameters described in the previous Section reflect

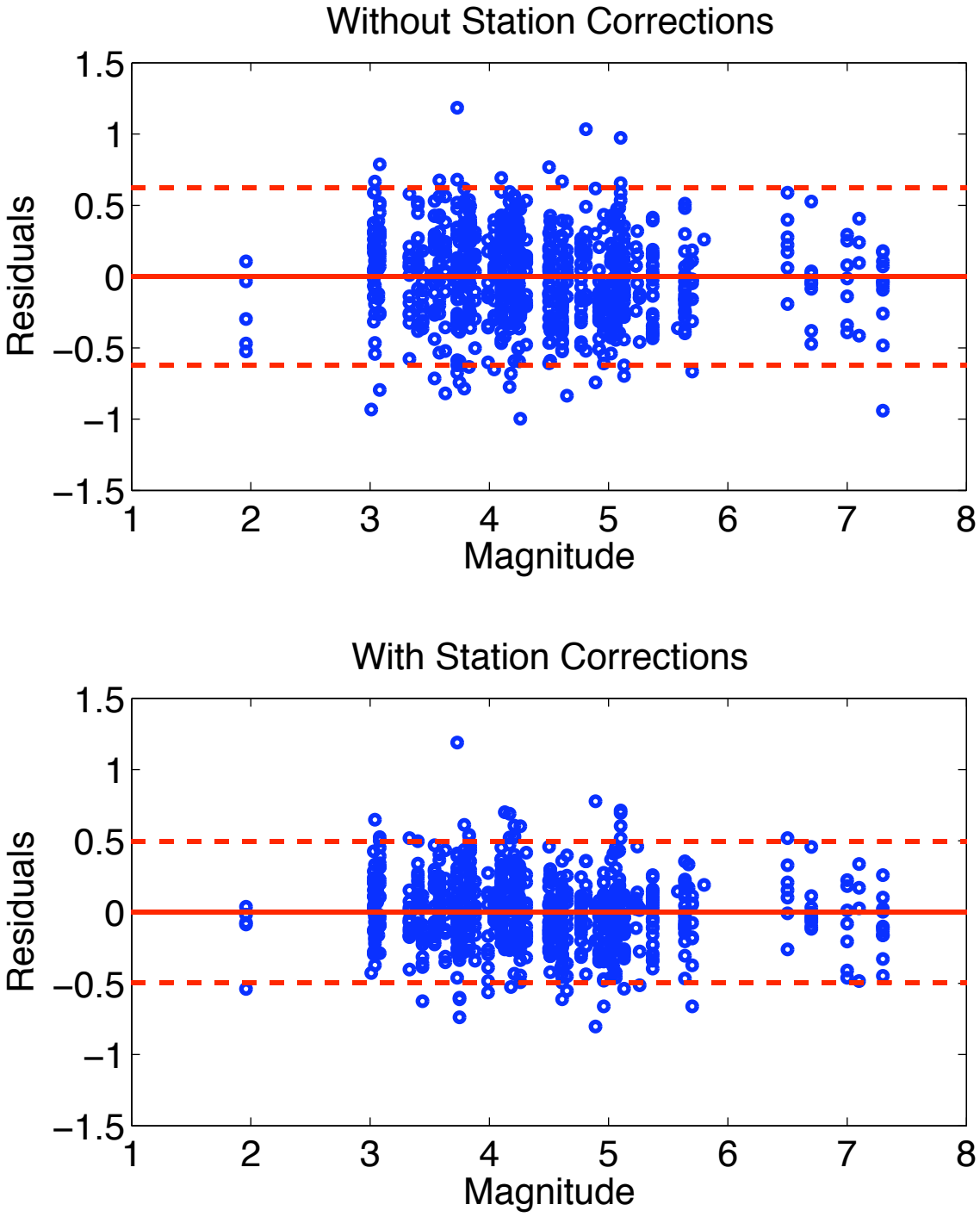

Figure 2.13: Residuals, $\epsilon = log_{10}(A_{obs}) - log_{10}(A_{pred})$, where A is rms S-wave acceleration amplitudes recorded on rock sites, plotted against magnitude. There are no systematic trends in the residuals with respect to magnitude.

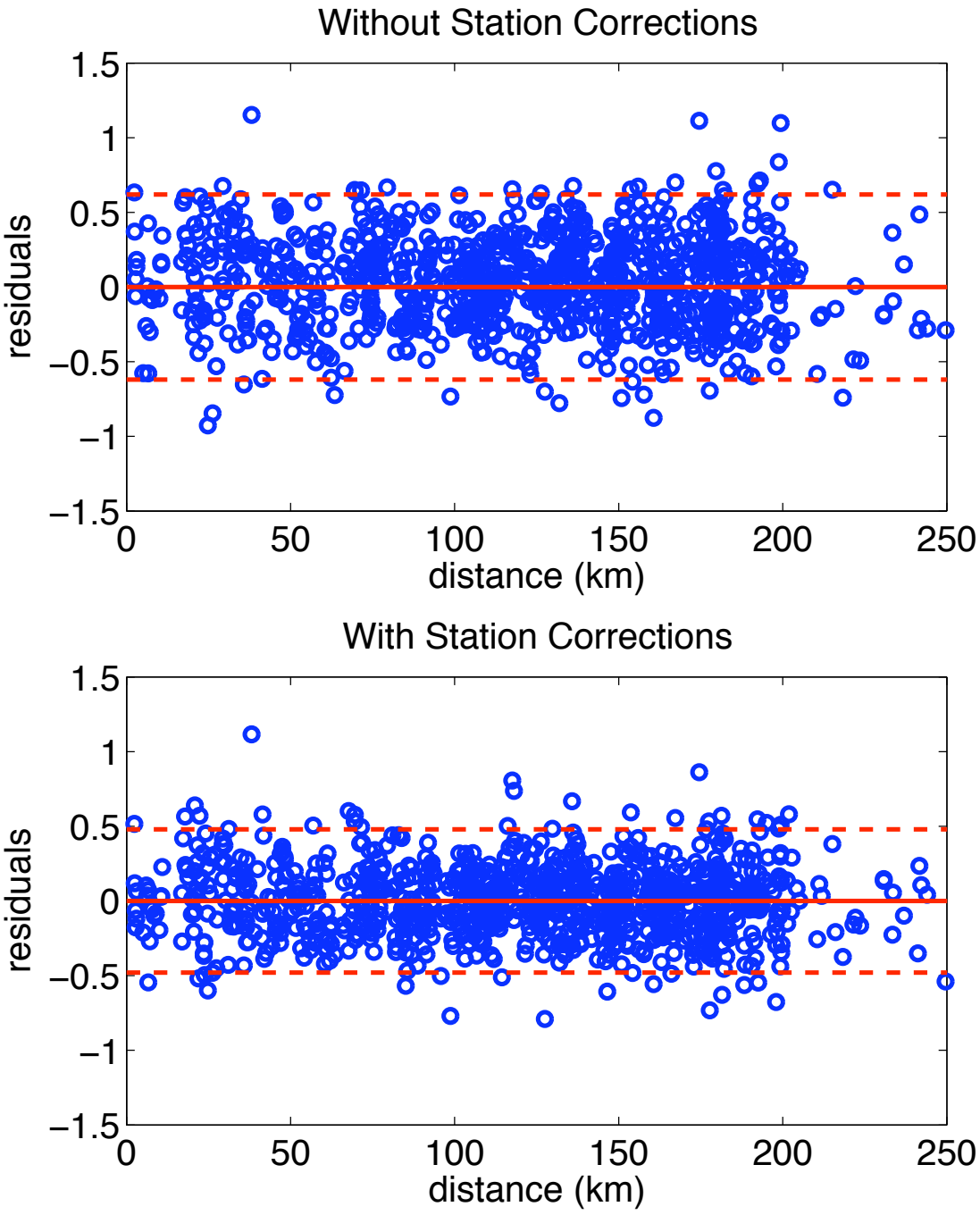

Figure 2.14: Residuals, $\epsilon = log_{10}(A_{obs}) - log_{10}(A_{pred})$, where A is rms S-wave acceleration amplitudes recorded on rock sites, plotted against distance. There are no systematic trends in the residuals with respect to distance.

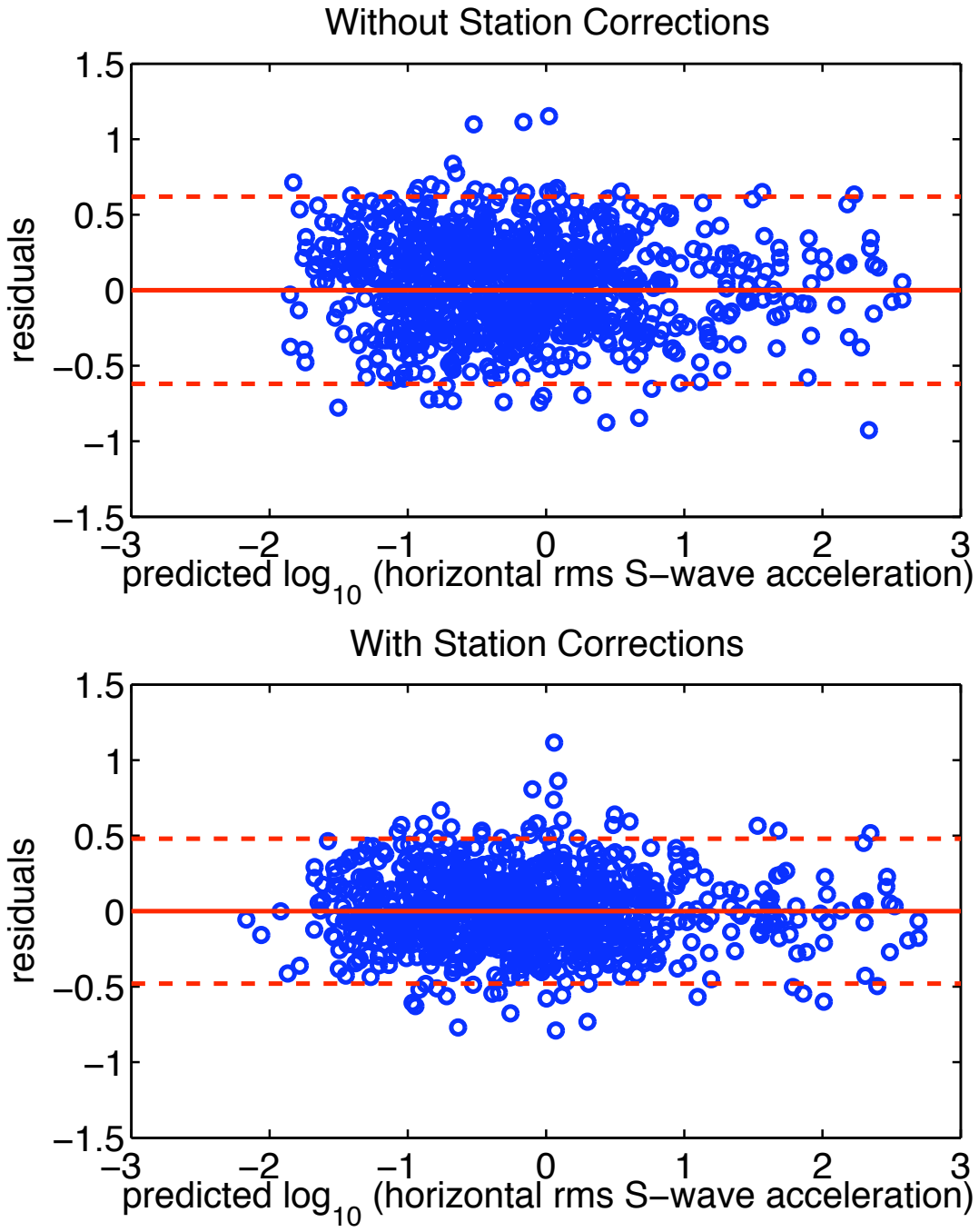

Figure 2.15: Residuals,$\epsilon = log_{10}(A_{obs}) - log_{10}(A_{pred})$, where A is rms S-wave acceleration amplitudes recorded on rock sites, plotted against predicted values. There are no systematic trends in the residuals relative to the predicted values.

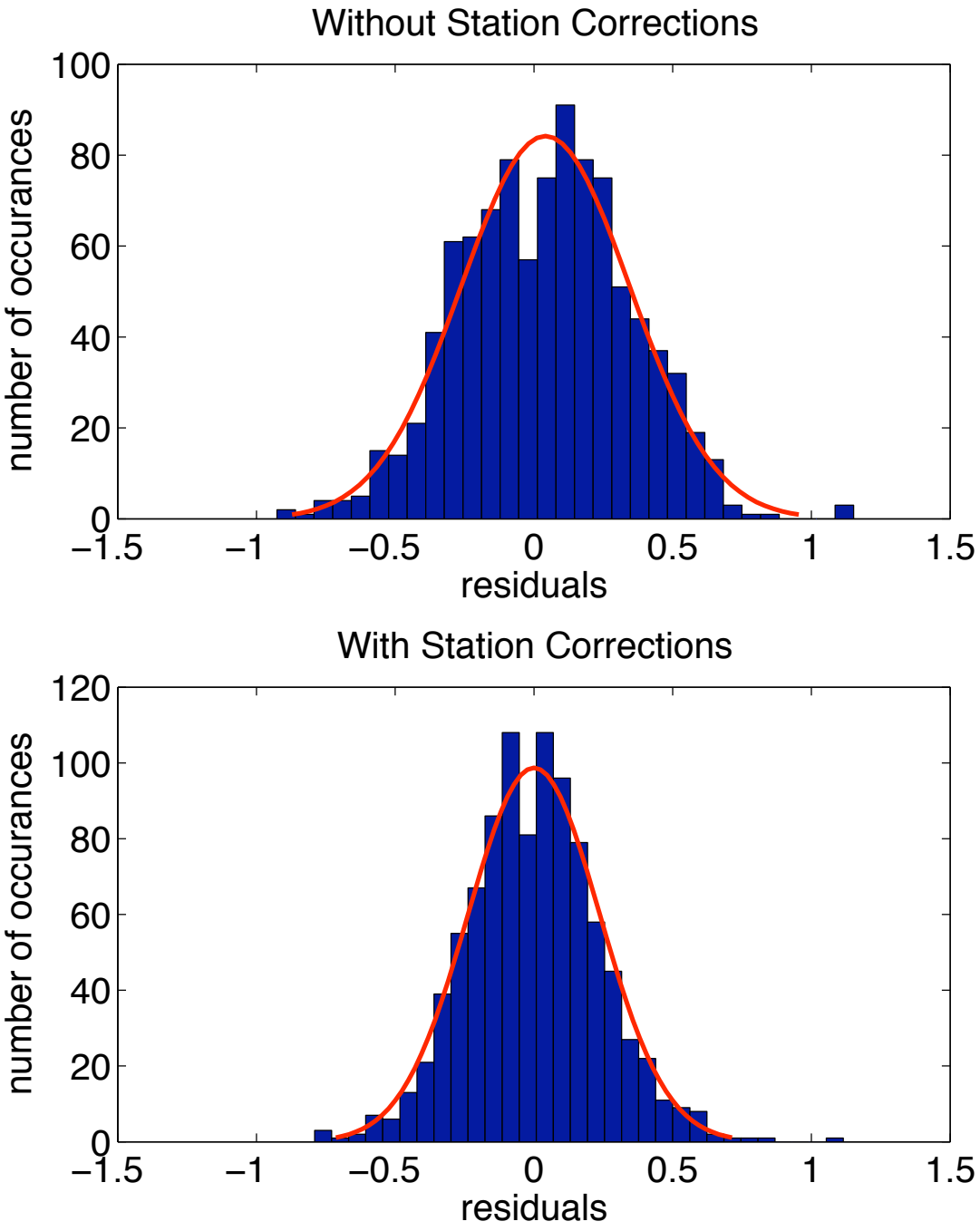

Figure 2.16: Histogram of residuals, $\epsilon = log_{10}(A_{obs}) - log_{10}(A_{pred})$, where A is rms S-wave acceleration amplitudes recorded on rock sites, with the best fit normal distribution shown for reference. The residuals satisfy the assumption that they are normally distributed about a mean of 0 with a constant variance, σ^2. The residuals for the station-corrected case have a smaller spread about 0, indicating the reduction in variance from the station corrections.

the influences of Boore, Joyner, and Fumal (1994, 1997) and Campbell (1981, 1997, 2002). However, to predict envelopes of ground motion (as shown in Figure 2.5), attenuation relationships are required for the remaining envelope parameters: rise times, durations, and 2 decay parameters for P- and S-waves. There is not much in the literature to guide the development of ground motion models for these envelope parameters.

Let *param* denote the parameters rise time, duration, and decay parameters τ and γ. The approach to modeling each element of *param* as a function of magnitude and distance is as follows:

- model $\log_{10}(param)$ (instead of linear *param*), since rise times, durations, and decay parameters are all positive quantities
- model the $\log_{10}(param)$ as a simple function of magnitude and distance

$$\log_{10}(param) = aM + bR + c\log_{10} R + d + \epsilon \tag{2.9}$$

  with M, R, and ϵ are as defined in Eqn. 2.3; $\epsilon \sim NID(0, \sigma^2)$ is assumed
- determine the statistically significant parameters
- identify and address trade-offs with other parameters

Figure 2.17 shows the various envelope parameters used to describe the evolution of the S-wave amplitudes with time according to Eqn. 2.2.

2.4.1 S-wave rise time, t_{rise_S}

The attenuation relationship obtained for S-wave rise time, t_{rise_S} is

$$\begin{aligned} \log_{10}(t_{rise_S}) &= 0.64M + 0.48\log_{10}(R) - 0.89 \\ \sigma_{t_{rise_S}} &= 0.23 \end{aligned} \tag{2.10}$$

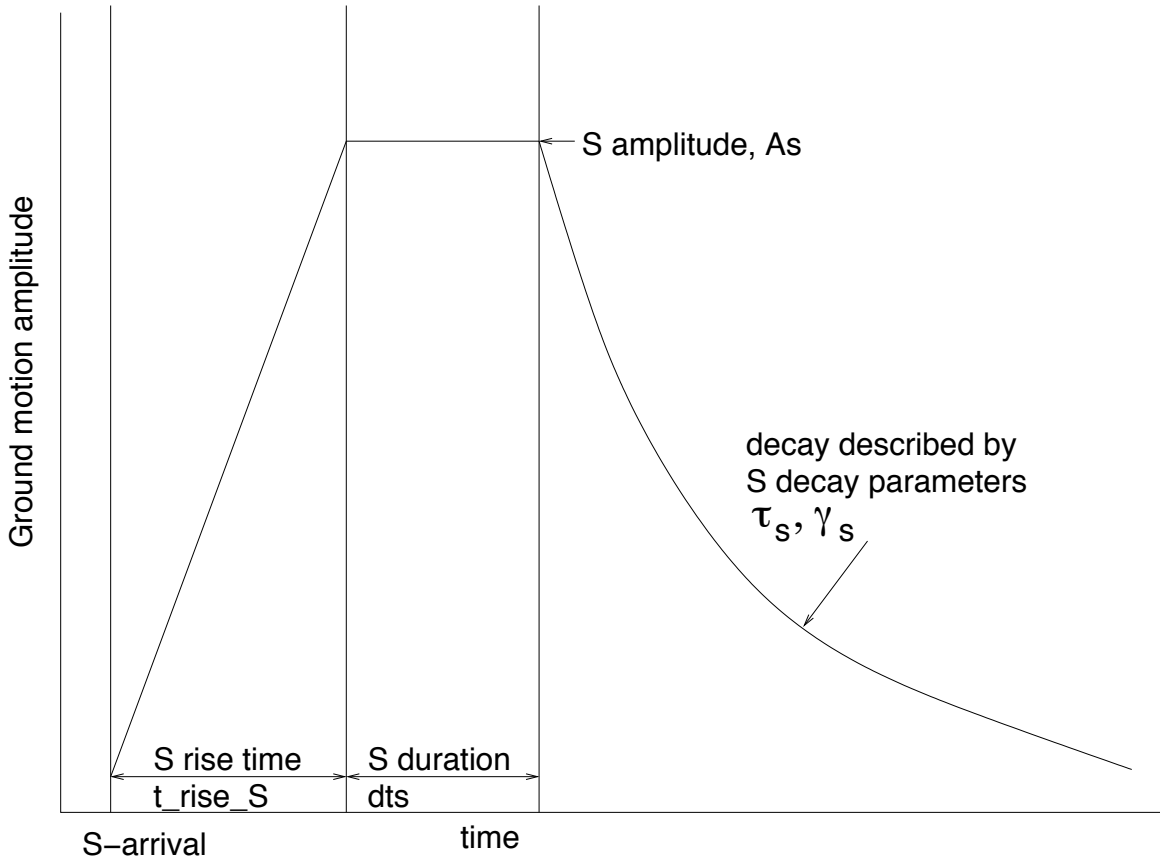

Figure 2.17: The various envelope parameters used to describe the evolution of S-wave amplitudes with time, according to Eqn. 2.2. The S-wave envelope parameters are rise time (t_{rise_S}), a constant amplitude (A_S), a duration (Δt_S), and two decay parameters (τ_S, γ_S).

This implies that the S-wave rise time, t_{rise_S} scales with seismic moment M_o as follows:

$$\begin{aligned} t_{rise_S} &\propto 10^{.64M}\sqrt{R} && (2.11)\\ &\propto 10^{.64\frac{2}{3}\log M_o}\sqrt{R} && (2.12)\\ &\propto M_o^{\frac{4}{9}}\sqrt{R} && (2.13) \end{aligned}$$

Figure 2.18 plots t_{rise_S} (in $\log_{10}$ units) plotted against magnitude and distance, as well as the residuals (observed rise times minus those predicted by Eqn 2.11) against the rise times predicted by Eqn. 2.11. A histogram of the residuals indicates that the assumption of normally and independently distributed errors is valid.

2.4.2 S-wave duration, Δt_S

The attenuation relationship obtained for S-wave duration, Δt_S, is

$$\begin{aligned} \log_{10}(\Delta t_S) &= -4.9\times 10^{-4}R - 0.13\log_{10} R && (2.14)\\ \sigma_{\Delta t_S} &= 0.21 \end{aligned}$$

Magnitude is not a statistically significant predictor of Δt_S. From plots of S-wave duration against magnitude and distance in Figure 2.19, neither M nor R are good predictors. The inadequacy of Eqn. 2.14 as a model for Δt_S is evident from the plot of residuals against predicted values. The presence of structure in the residual plot, in particular, the increasing scatter with increasing $\log_{10}(\text{duration}_S)$, and a histogram of the residuals (observed duration minus those predicted by Eqn. 2.14) indicates that a normal distribution is not appropriate and perhaps different predictors are necessary. It is possible that the S-wave duration Δt_S is trading off with one of the decay parameters, τ_S, which is also a time quantity.

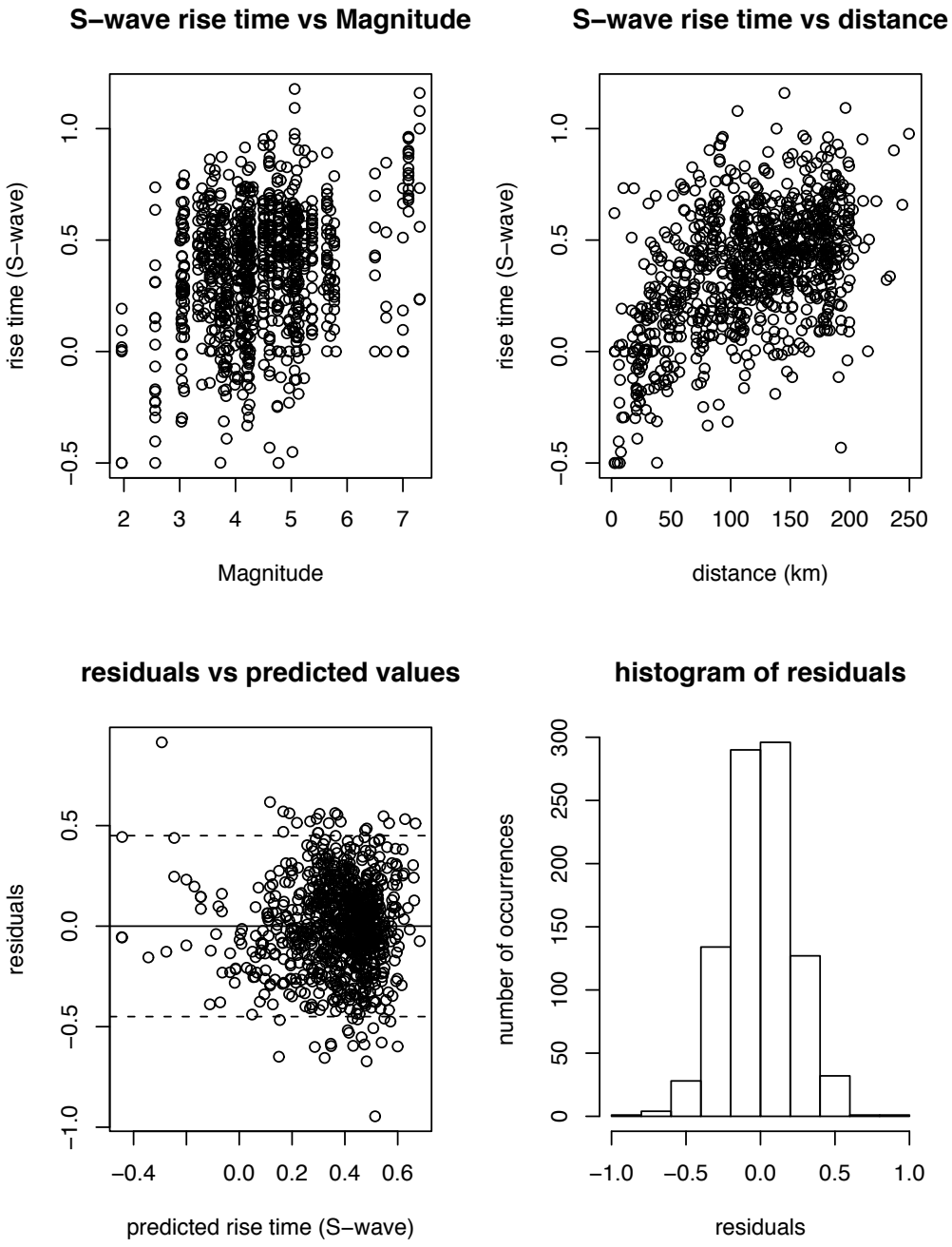

Figure 2.18: S-wave rise time, t_{rise_S}, plotted against magnitude and distance. Some regression diagnostic plots: residuals (observed rise times minus those predicted by Eqn. 2.11) against predicted values, and histogram of residuals.

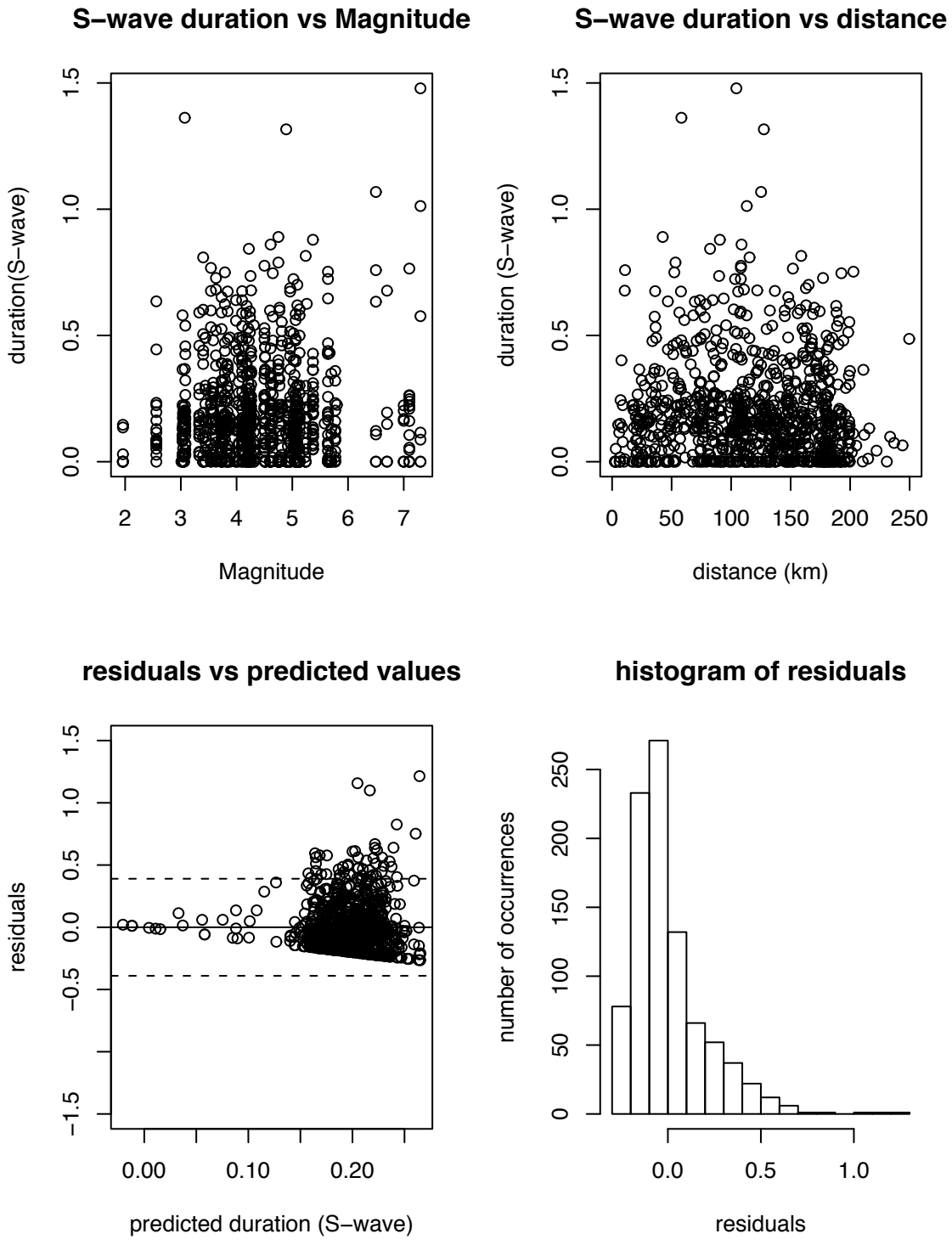

Figure 2.19: S-wave duration, Δt_S plotted against magnitude, and distance. Some regression diagnostic plots: residuals (observed S-wave duration minus those predicted by Eqn. 2.14) against predicted values and histogram of residuals.

2.4.3 S-wave decay parameter τ_S

The attenuation relationship obtained for S-wave decay parameter τ_S is

$$\begin{aligned} \log_{10}(\tau_S) &= 0.04M + 0.39\log_{10} R + 1.73\gamma_S - 0.59 \\ \sigma_{\tau_S} &= 0.18 \end{aligned} \tag{2.15}$$

The parameterization stated in Eqn. 2.2 suffers from trade-offs between the two decay parameters, τ, γ. The strongest predictor of τ is the other S-wave decay parameter γ; magnitude and distance are weaker, but still statistically significant predictors of τ. τ_S is also probably trading off with the S-wave duration, Δt_S. Figure 2.20 show τ_S plotted against magnitude, distance, and decay parameter γ_S. Plots of residuals (observed τ_S minus the predictions from Eqn. 2.15) against predicted values and a histogram of residuals relative to Eqn 2.15 indicate that $\log_{10}(\tau_S)$ is indeed normally distributed, and that the assumption of normally and independently distributed residuals is valid.

2.4.4 S-wave decay parameter γ_S

The attenuation relationship for S-wave decay parameter γ_S is

$$\begin{aligned} \log_{10}(\gamma_S) &= -0.014M - 5.28\times10^{-4}R - 0.11\log_{10} R + 0.38\tau_S + 0.26 \\ \sigma_{\gamma_S} &= 0.09 \\ \text{average, } \bar{\gamma_S} &= 0.15 \end{aligned} \tag{2.16}$$

Decay parameter τ_S is the strongest predictor of γ_S; magnitude and distance are statistically significant, but weaker predictors. The model given in Eqn. 2.16 appears adequate from examining plots of residuals against predicted values and a histogram of residuals.

As mentioned earlier, there are trade-offs between τ_S and γ_S. Eqn. 2.15 and 2.16 cannot both be used; it is a circular relationship. Since γ has a smaller range

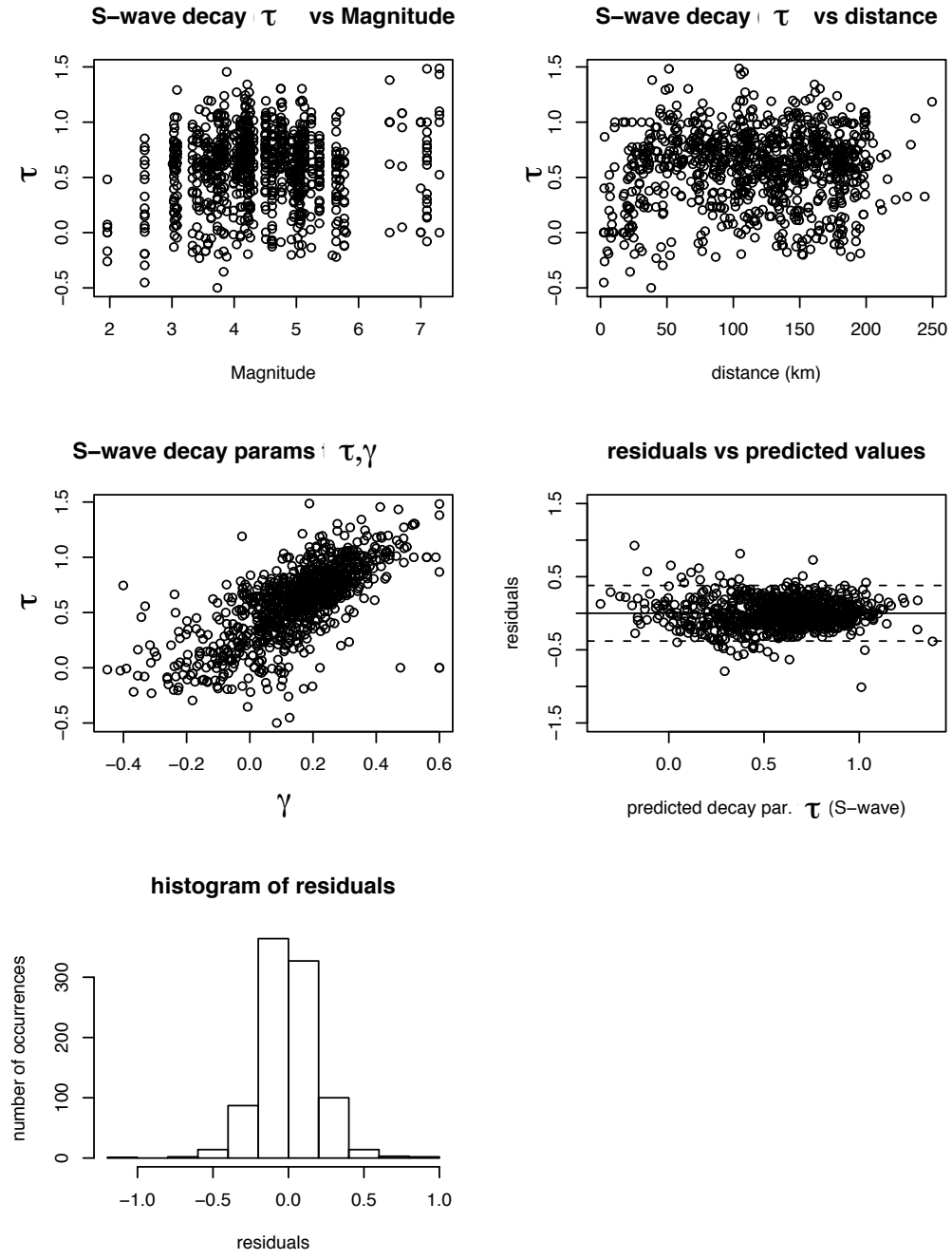

Figure 2.20: S-wave decay parameter τ_S plotted against magnitude, distance, S-wave decay parameter γ_S. Some regression diagnostic plots: residuals against predicted values and histogram of residuals.

$(-0.4, 0.6)$ than $\tau(-0.5, 1.5)$, a constant average value, $\bar{\gamma_S}$, is used as the model for γ_S, and Eqn. 2.15 is used to model τ.

2.5 Summary

In this chapter, an approach to characterizing ground motion envelopes as functions of magnitude, distance, and site was described. The dataset used was discussed and a new parameterization of ground motion envelopes that decomposes the observed envelope into P-wave, S-wave, and ambient noise envelopes was introduced. This characterization scheme (Eqns. 2.1 and 2.2) uses 11 parameters (rise time, duration, amplitude, and 2 decay parameters for each of the P- and S-waves, and a constant for ambient noise) to describe a particular observed ground motion envelope. The Neighbourhood Algorithm (Appendix B) is used to find the set of envelope parameters that minimize the least squares error for each envelope history in the database. In an effort to better resolve trade-offs between the various envelope parameters, two iterations on fitting envelope parameters to the entire database and developing envelope attenuation relationships were performed. After the first iteration, envelope attenuation relationships describing magnitude, distance, and site dependence of the 11 envelope parameters were obtained. The P- and S-wave amplitude attenuation relationships reflect influences of traditional strong motion attenuation relationships (Boore and Joyner, 1982; Boore et al., 1993; Boore et al., 1997; Campbell, 1981). Simple linear regression is used for the rise time, duration, and decay parameters. These preliminary attenuation relationships were used to predict more informed starting models for the second iteration of fitting envelopes. (Like most nonlinear problems, fitting envelopes with this parameterization is somewhat sensitive to the initial models. Using the Neighbourhood Algorithm makes this problem less sensitive to initial models.) The envelope attenuation relationships that will be used in the subsequent chapters are based on the second iteration of envelope fitting.

The root mean square horizontal S-wave acceleration envelope parameters are used to illustrate the process of developing predictive relationships for rise time, amplitude,

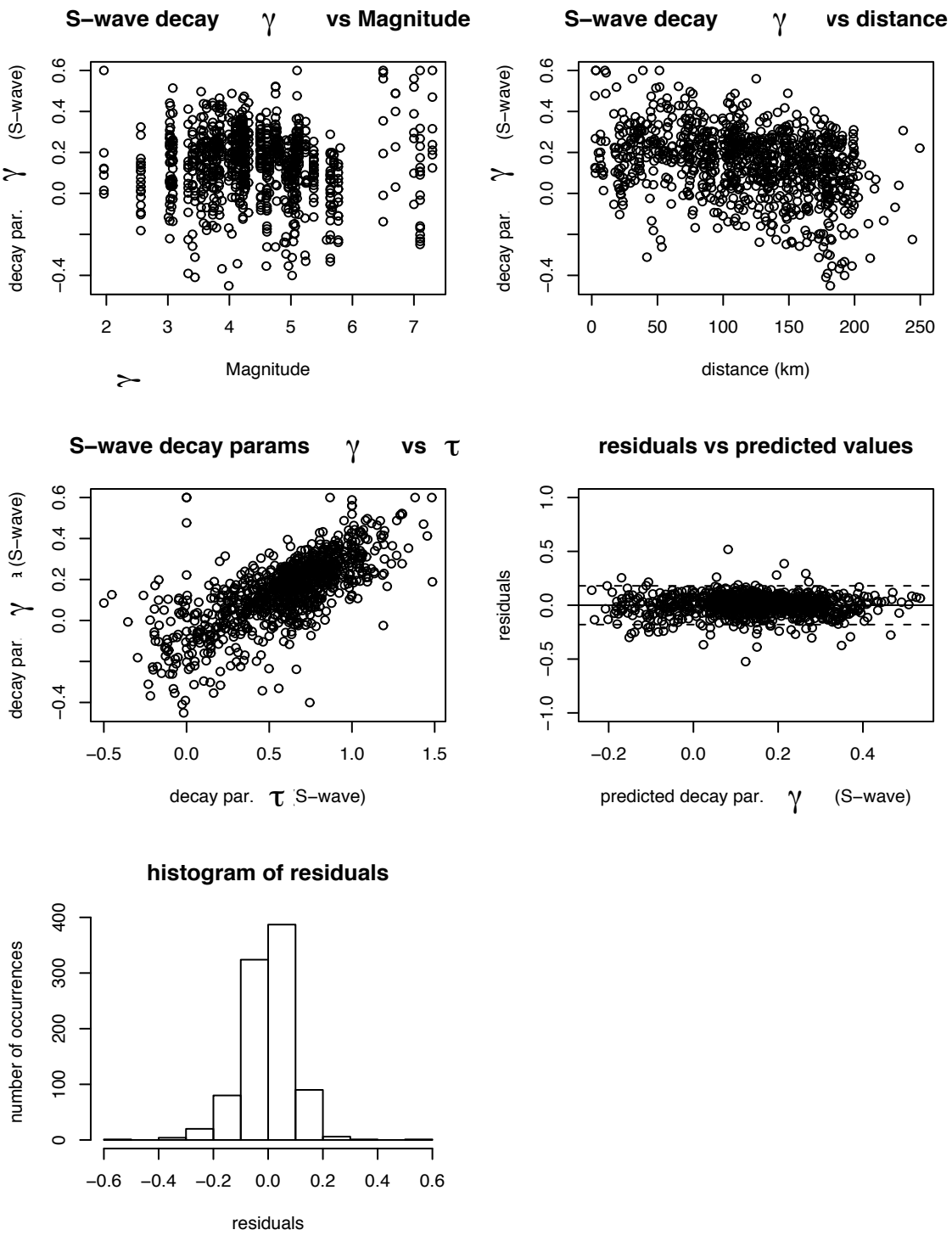

Figure 2.21: S-wave decay parameter γ_S plotted against magnitude, distance, and S-wave decay parameter τ_S. Some regression diagnostic plots: residuals against predicted values and histogram of residuals.

duration, and decay parameters τ and γ. Identical processes are used to develop envelope attenuation relationships for the P-waves, and for the other channels of ground motion. Thus, given the envelope attenuation relationships, and a particular magnitude, distance, and site, 11 envelope parameters can be predicted. Given 11 envelope parameters, Eqn. 2.2 can then be used to generate the envelope as a function of time.

One may ask: were there any gains in iterating on the envelope fitting procedure? Included in Table 2.2 are the first and second iterations for the horizontal S-wave (rock sites) envelope attenuation relationships. The standard errors, σ, for the various envelope parameters are listed in the last column. In general, the second iteration shows a slight decrease in the standard errors, indicating some improvement over the first set of results. The improvement was not large enough to motivate additional iterations. The envelope-fitting procedure is computationally expensive. It takes approximately 336 hours on 6 processors to complete one iteration of fitting envelopes to the entire database.

Iteration 1

	M	R	$\log_{10} R$	c_1	c_2	τ	γ	intercept	σ
$T_{rise,P}$	0.06	-	0.31	-	-	-	-	-0.47	0.24
A_P	0.75	3.52×10^{-3}	1.07	1.95	1.09	-	-	-1.43	0.42
ΔT_P	0.023	-	0.38	-	-	-	-	-0.34	0.31
τ_P	0.029	1.85×10^{-3}	0.23	-	-	-	-	-0.26	0.40
γ_P	-0.030	-2.0×10^{-3}	-	-	-	-	-	0.45	0.22
$T_{rise,S}$	0.058	-	0.31	-	-	-	-	-0.48	0.24
A_S	0.80	2.97×10^{-3}	1.34	1.66	1.18	-	-	-0.70	0.32
ΔT_S	0.030	-1.19×10^{-3}	0.18	-	-	-	-	-0.21	0.21
τ_S	0.050	-	0.42	-	-	-	1.18	0.52	0.22
γ_S	-0.026	-	-0.22	-	-	-	-	0.52	0.20
noise								-2.47	

Iteration 2

	M	R	$\log_{10} R$	c_1	c_2	τ	γ	intercept	σ
$T_{rise,P}$	0.06	5.5×10^{-4}	0.27	-	-	-	-	-0.37	0.22
A_P	0.72	3.3×10^{-3}	1.20	1.6	1.05	-	-	-1.06	0.31
ΔT_P	-	2.58×10^{-3}	0.21	-	-	-	-	-0.22	0.39
τ_P	0.047	-	0.48	-	-	-	0.82	-0.75	0.28
γ_P	-0.032	-1.81×10^{-3}	-0.1	-	-	0.27	-	0.64	0.16
$T_{rise,S}$	0.64	-	0.48	-	-	-	-	-0.89	0.23
A_S	0.78	2.6×10^{-3}	1.35	1.48	1.11	-	-	-0.64	0.31
ΔT_S	-	-4.87×10^{-4}	0.13	-	-	-	-	0.0024	0.21
τ_S	0.037	-	0.39	-	-	-	1.73	-0.59	0.18
γ_S	-0.014	-5.28×10^{-4}	-0.11	-	-	0.38	-	0.26	0.09
noise								-2.0	

Table 2.2: Envelope attenuation relationships for rms horizontal acceleration on rock sites: iterations 1 and 2.