

Chapter 4

Data Processing

4.1 Introduction

In order to have a robust unbiased estimate of the occurrence or non-occurrence of liquefaction it is of preeminent importance to have the highest quality data. A probabilistic correlation requires powerful statistical techniques, but is only as good as the quality of data to which the techniques are applied. To this end, data processing was of utmost importance in this study. A considerable amount of time was spent processing and reviewing the database to minimize epistemic uncertainty that can creep in due to human error, biased interpretation, and poor analysis techniques.

4.2 Field Observations

The basis of a liquefaction correlation is a research engineer's observation of liquefaction or absence of liquefaction following a seismic event, and the index test measurements of the suspect critical layer. This basis is inherently fraught with uncertainty including; lack of full coverage of affected area, misinterpretation of field evidence, poor index testing procedures, difficult field conditions, etc.

One of the primary discrepancies of a database of this type is that researchers tend to retrieve more liquefied than non-liquefied case histories. This can be attributed to the fact that testing in a liquefied area is much more appealing than testing at a site that

hasn't experienced liquefaction. This unfortunately leads to a data bias, more liquefied case histories than non-liquefied case histories. To account for this data imbalance the procedure of bias weighting is used, as described in Chapter 5 on Bayesian analysis.

Liquefaction field correlations is that they are not based truly on the occurrence or non-occurrence of liquefaction but on the observation of the manifestations of liquefaction at a particular location and the lack of manifestation at some other location. These manifestations can take the form of sand boils or sand blows, lateral spreading, building tilting or settlement, ground loss, broken lifelines, etc. Liquefaction can and does occur at depths where there is no surface evidence of the event. This of course does not make it into a liquefaction database, it fits the tree-falling-in-the-woods analogy.

The most content rich sites are sites that are labeled as marginal. Marginal liquefaction does not exist, a soil deposit either liquefies or does not liquefy. Marginal is a research engineer's interpretation that at this location liquefaction was either incipient or occurred and resulted in minimal surface manifestations. These sites are included in the database and tend to have the most information content because they fall near the limit-state (threshold of liquefaction/non-liquefaction).

All these vagaries are incorporated into the database and can result in epistemic uncertainty. To minimize this uncertainty a panel of experts reviewed the database and came to a consensus on each site and the data it contained. This process of consensus

results in a robust database that contains the best assessment of each variable to the highest standards of practice.

4.3 Strength Parameters

4.31 Choice of Logs

At any given site there can be multiple CPT and also corollary SPT logs to choose from. Proximity of the logs to the observed liquefaction/non-liquefaction is critical. The depositional environment and the properties that lead to liquefaction can vary significantly over a small distance and therefore it is important to be as close to the observed location as possible. Logs that are considered to be representative of the conditions are chosen. When there are multiple logs, the values (such as tip and sleeve resistance) are average.

CPT logs that were measured using a mechanical cone or a sleeveless cone are not used in this database because of the lack of sleeve measurements. However when a sleeveless cone trace has a corollary SPT log that shows that the critical layer is composed of clean sand ($FC < 5\%$) then the tip resistance is used in conjunction with a prescribed median “clean sand” friction ratio ($R_f \cong 0.35\%$). This allows the use of important early CPT tip resistance data with a neutral friction ratio.

There are a few earthquake reconnaissance trips that utilized a Chinese cone. The report by Earth Technology (1985) showed that there is very little difference between tip and

sleeve readings using the Chinese cone and a cone following ASTM specifications (D3441 & D5778). Therefore the Chinese cone was treated no different in this database.

4.3.2 Case Selection

The objective in case selection in this study was to end up with a group of statistically independent data points. Some previous correlations have used multiple liquefaction or non-liquefaction cases from a single site to generate more statistical data for analysis. This method can be incorrect for two reasons. First, given a site with consistent stratigraphy and a uniform depositional environment, selecting two liquefied or two non-liquefied cases from the same critical layer results in cross-correlation of these two data points. The cross-correlation must be accounted for in any form of statistical analysis and will result in much higher uncertainty or much reduced informational content for each data point. Second, if a particular layer within the site does liquefy this then modifies the incoming seismic energy for the layers above through seismic isolation and below by blocking full reflection off the surface. This leads to a modified CSR for other layers at the site which can be difficult to determine.

4.3.3 Critical Layer Selection

Selection of the critical layer is an important step in estimating the seismic strength of a particular soil deposit. The criteria for selection is finding the strata of soil that is the weakest-link-in-the-chain from a liquefaction perspective. Finding the weakest link requires observing the tip resistance and friction ratio in conjunction, with the addition of

a SPT log, for soil classification, if one is available. For most depositional environments this can be a simple matter of looking for the smallest continuous stretch of tip resistance with low friction ratio that agrees with the SPT log in terms of a liquefiable material. It can be a difficult proposal for fluvial depositional environments where the strata are thin, interbedded, and discontinuous both horizontally and vertically. A final criteria for identifying a critical layer is comparing the suspect layers to previous correlations. This aids in the more difficult sites where determining which of multiple layers liquefied or didn't liquefy.

One issue that is not commonly addressed in liquefaction correlations is that the *in situ* data is usually acquired post ground shaking. Particularly for the liquefied cases, the soil strength and properties have most likely been modified due to the process of liquefaction. Chameau et al. (1991) looked at sites that were affected by the Loma Prieta Earthquake in which previous CPT data existed. Post event CPT data was acquired and compared to the pre-event CPT data. They found that loose materials experienced the most alteration in tip resistance due to the ground shaking and subsequent liquefaction. This comes as no surprise. Recent work involving large scale liquefaction blast tests have and are being performed in Japan where pre- and post-liquefaction CPT measurements are made. Hopefully this data will resolve the bias and allow for proper accounting of the changes that occur within a liquefied layer.

If it can be assumed that tip resistance has a positive correlation with relative density for clean sands (Schmertmann, 1978), then the greater the tip resistance the greater the

relative density. In a critical state framework, given a constant confining stress, the higher the relative density (lower the void ratio) the less capacity the soil has for contractive behavior. Liquefaction is premised on this contractive behavior of soils. Therefore, the closer a point lies to the limit-state or liquefaction boundary the less contractive it is, and the less pre- to post-liquefaction change in resistance it is likely to experience. On the non-liquefaction side of the limit-state it is assumed that the resistance is unmodified by the ground shaking because no liquefaction has occurred. Another issue is that if a CSR value is determined for a liquefied site using the post-liquefaction *in situ* measurements for site response analysis, the value may be slightly higher than pre-liquefaction conditions because of the stiffening that has occurred.

Given all these pre- and post-liquefaction considerations, it is conjectured that the limit-state function is totally unaffected by post-liquefaction densification because:

1. near the limit-state the soils are near the critical state (small state parameter) and therefore have not significantly densified,
2. non-liquefied soils will have no post-event densification and therefore are unaffected by the event and will maintain their position near the limit-state.

The soils most affected by liquefaction, that will give vastly different post-event resistance measurements are the loose or low tip resistance soils, and these have little impact on the limit-state function in a Bayesian-type analysis.

4.3.4 Index Measurements

Once the critical layer has been selected it is a matter of determining the appropriate statistics of the measurements within the layer. Kulansingam, Boulanger, & Idriss (1999) studied various procedures for estimating an average tip resistance over a standardized distance of cone travel. They looked at different standardized distances and came to the conclusion that having a preset distance over which the resistance is averaged works poorly.

The approach used in this study was to let the depositional environment dictate. Using the procedures described above for identifying the critical layer, the maximum distance over which the soil deposit lies is often apparent. The top and bottom depths are taken as extrema. The distribution of the tip and sleeve resistances are assumed to be normal, and the averages and standard deviations are calculated from a digitized form of the trace. Raw sleeve and tip measurements are used to calculate the friction ratio in order to eliminate aliasing that may have occurred in the field calculations.

Induced pore pressure can have an affect on the tip and sleeve measurements. This affect is pronounced in soils that respond in an undrained manner to the strain imposed by the advancing cone (i.e. fine grained soils). For most soils that are susceptible to liquefaction, fully drained cone penetration is assumed (Lunne et al., 1997). Therefore, in general, no pore pressure corrections are necessary for materials that are potentially liquefiable. The assumption of fully drained response was checked using pore pressure measurements, when available, for each site.

4.3.5 Masked Liquefaction

In certain situations liquefaction may occur at depth but evidence may not reach the ground surface due to the monolithic or unified nature of overlying non-liquefiable strata. This masked liquefaction situation was researched and presented by Ishihara (1985). The results from that research are used to screen sites that are found to be liquefiable in terms of the index measurements, has overlying non-liquefiable material that fits the Ishihara (1985) thickness criteria, showed no surface manifestation of liquefaction, and was reported as a non-liquefied site. For reference, at a site experiencing a low level of ground shaking ($PGA < 0.2 \text{ g}$) with a 2 m thick liquefiable layer, an overlying non-liquefiable layer of approximately 2 m could eliminate all surface manifestation of liquefaction.

4.3.6 Screening for Other Failure Mechanisms

Certain soil types are not susceptible to liquefaction but may deform via cyclic softening. These soils may exhibit surface manifestations that can appear quite similar to what may be observed in “classic” liquefaction, such as lateral spreading, and building tilting, punching, and settlement. However it has been shown that the failure mechanism is quite different from liquefaction and is primarily a function asymmetrical driving shear stresses (K_α). The soils that are susceptible to cyclic softening tend to have a high percentage of fines and these fines will tend to behave in a plastic manner. Several cases like this were observed in the 2001 Kocaeli, Turkey Earthquake and the 2001 Chi-Chi, Taiwan Earthquake. Since the limit-state and the overall correlation is based on “classic” liquefaction, it is not appropriate to include these cases in the analysis.

A criteria for screening these cases is based on research of fines content and plasticity in relation to liquefaction susceptibility (Andrews & Matin, 2000; Andrianopoulos et al., 2001; Guo & Prakash, 1999; Perlea, 2000; Polito, 2001; Sancio et al., 2003; Yamamuro & Lade, 1998, Youd & Gilstrap, 1999; to name a few). The criteria for soils not susceptible to liquefaction used in this study are shown graphically in Figure 4.1.

4.3.7 Normalization

The tip and sleeve are normalized using the variable normalization scheme presented with this study in Chapter 3, on Normalization. Note that the tip and sleeve values are normalized equivalently which results in no change for a normalized friction ratio ($R_{f,1} = R_f$).

4.3.8 Thin Layer Correction

Thin layer corrections, if they were required, are performed using the method proposed in this study in Chapter 2, on Thin Layer Correction. Note that only 4% of the cases in the database required a thin layer correction. For database purposes the thin layer correction was limited to a maximum of 1.5 ($C_{thin} \leq 1.5$).

4.3.9 Processed Strength Parameters

The result of this processing procedure is unbiased, statistically independent $q_{c,1}$, $f_{s,1}$, and R_f values for the liquefied and non-liquefied cases. These are mean resistance values and variances over the extent of the critical layer which have been normalized to one atmosphere and corrected for thin layer issue if required.

4.4 Stress Parameters

4.4.1 Cyclic Stress Ratio

The dynamic stress that the critical layer experienced is determined using the simplified uniform cyclic stress ratio as defined by Seed & Idriss (1971),

$$CSR = 0.65 \cdot \frac{a_{\max}}{g} \cdot \frac{\sigma_v}{\sigma'_v} \cdot r_d \quad (4.1)$$

The CSR value calculated using Equation 4.1 is assumed to be the average or mean of a normally distributed random variable as in Equation 4.2. The variance of CSR is calculated via equation 4.3, where the coefficient of variation is equal to the standard deviation divided by the mean. Both Equation 4.2 and 4.3 are using first-order Taylor series expansions about the mean point, including only the first two terms.

$$\mu_{CSR} \cong 0.65 \cdot \frac{\mu_{a_{\max}}}{g} \cdot \frac{\mu_{\sigma_v}}{\mu_{\sigma'_v}} \cdot \mu_{r_d} \quad (4.2)$$

$$\delta_{CSR}^2 \cong \delta_{\alpha_{max}}^2 + \delta_{rd}^2 + \delta_{\sigma_v}^2 + \delta_{\sigma'_v}^2 - 2 \cdot \rho_{\sigma_v \sigma'_v} \cdot \delta_{\sigma_v} \cdot \delta_{\sigma'_v} \quad (4.3)$$

Total and effective stress are correlated parameters, therefore the inclusion of the correlation coefficient term for these two variables is necessary.

4.4.2 Peak Ground Acceleration

The geometric mean of the peak ground acceleration is based on the best estimation of ground shaking possible. The methods of estimation are; strong motion recordings, site response, calibrated attenuation relationships, adjustment of estimated site pga through general site response modeling, and general attenuation relationships. Using a calibrated attenuation relationship means using all available recordings to tune general attenuation relationships for event-specific variations and azimuth specifics when recordings permit.

The coefficient of variation of the peak ground acceleration is fixed according to the method of ground shaking estimation;

- $\delta < 0.10$ for sites with strong motion stations less than 100m from site,
- $\delta = 0.10$ to 0.25 for sites with strong motion stations within 100 to 500m from site or where site response analysis was performed using a nearby rock recording as input base motion ,
- $\delta = 0.25$ to 0.35 for sites with strong motion stations within 500m to 1000m and/or estimates from calibrated attenuation relationships,
- $\delta = 0.35$ to 0.5 for others.

This is a subjective determination of the variance of the ground shaking but is based on typical uncertainty bands from general attenuation relationships that have coefficient of variations of between 0.3 and 0.5 (e.g. Abrahamson & Silva, 1997).

4.4.3 Total and Effective Stress

The total and effective vertical stress are correlated variables and this correlation must be accounted. The critical layer is selected using the procedures outlined above. From this the total extent of the critical layer is used to calculate the mean and variance of the critical layer, assuming that it is normally distributed. The variance is estimated using a 6 sigma approach, where the extrema of the layer are assumed to be three standard deviations away from the mean on either side. The total variance is then divided by six to give an estimate of the standard deviation.

A deterministic estimate is made of the mean unit weight of the soil above and below the water table. The variance is based on statistical studies of the measured variability of soil unit weight and is set at $\delta \cong 0.1$ (Kulhawy & Trautman, 1996). The water table mean is taken as the reported field measured value (with consideration given for the depth of water table during the seismic event) and the variance is set at a fixed standard deviation of $\sigma = 0.3$ m., a reasonable estimate of water table fluctuations given relatively stable groundwater conditions. An estimate of the total and effective vertical stresses, their respective variances, and covariance can then be calculated using the expansion Equations 4.4 through 4.9,

$$\mu_{\sigma_v} \cong \mu_{\gamma_1} \cdot \mu_{h_w} + \mu_{\gamma_2} \cdot (\mu_h - \mu_{h_w}) \quad (4.4)$$

$$\mu_{\sigma_v'} \cong \mu_{\gamma_1} \cdot \mu_{h_w} + (\mu_{\gamma_2} - \gamma_w) \cdot (\mu_h - \mu_{h_w}) \quad (4.5)$$

$$\sigma_{\sigma_v}^2 \cong \mu_{h_w}^2 \cdot \sigma_{\gamma_1}^2 + (\mu_h - \mu_{h_w})^2 \cdot \sigma_{\gamma_2}^2 + \mu_{\gamma_2}^2 \cdot \sigma_h^2 + (\mu_{\gamma_1} - \mu_{\gamma_2})^2 \cdot \sigma_{h_w}^2 \quad (4.6)$$

$$\sigma_{\sigma_v'}^2 \cong \mu_{h_w}^2 \cdot \sigma_{\gamma_1}^2 + (\mu_h - \mu_{h_w})^2 \cdot \sigma_{\gamma_2}^2 + (\mu_{\gamma_2} - \gamma_w)^2 \cdot \sigma_h^2 + (\mu_{\gamma_1} + \gamma_w - \mu_{\gamma_2})^2 \cdot \sigma_{h_w}^2 \quad (4.7)$$

$$Cov[\sigma_v, \sigma_v'] \cong (\mu_{h_w}^2 \cdot \sigma_{\gamma_1}^2) + (\mu_{\gamma_1} - \mu_{\gamma_2}) \cdot (\mu_{\gamma_1} + \gamma_w - \mu_{\gamma_2}) \cdot \sigma_{h_w}^2 + (\mu_h - \mu_{h_w})^2 \cdot \sigma_{\gamma_2}^2 + \mu_{\gamma_2} \cdot (\mu_{\gamma_2} - \gamma_w) \cdot \sigma_h^2 \quad (4.8)$$

$$\rho_{\sigma_v, \sigma_v'} = \frac{Cov[\sigma_v, \sigma_v']}{Var[\sigma_v] \cdot Var[\sigma_v']} \quad (4.9)$$

4.4.4 Nonlinear Shear Mass Participation Factor (r_d)

The nonlinear shear mass participation factor accounts for nonlinear response within a soil column and reduces the peak ground acceleration at the surface to reflect the ground acceleration that is experienced at the critical depth. This factor, denoted as r_d , has been derived from ground response analyses. In recent work, 2,153 site response analyses were run using 50 sites and 42 ground motions covering a comprehensive suite of motions and soil profiles (Cetin & Seed, 2000). This brute force approach allows for careful statistical analysis of the median response given the depth, peak ground

acceleration, moment magnitude, and indicative shear wave velocity of the site. The variance was estimated from the dispersion of these simulations. The median values can be estimated using Equations 4.10 and 4.11, and the variance from Equations 4.12 and 4.13,

$$\mathbf{d} < 65 \text{ ft} \quad (4.10)$$

$$r_d(d, M_w, a_{\max}) = \frac{\left[1 + \frac{-9.147 - 4.173 \cdot a_{\max} + 0.652 \cdot M_w}{10.567 + 0.089 \cdot e^{0.089 \cdot (-d - 7.760 \cdot a_{\max} + 78.576)}} \right]}{\left[1 + \frac{-9.147 - 4.173 \cdot PGA + 0.652 \cdot M_w}{10.567 + 0.089 \cdot e^{0.089 \cdot (-7.760 \cdot a_{\max} + 78.576)}} \right]} \pm \sigma_{\varepsilon r_d}$$

$$\mathbf{d} \geq 65 \text{ ft} \quad (4.11)$$

$$r_d(d, M_w, a_{\max}) = \frac{\left[1 + \frac{-9.147 - 4.173 \cdot a_{\max} + 0.652 \cdot M_w}{10.567 + 0.089 \cdot e^{0.089 \cdot (-d - 7.760 \cdot a_{\max} + 78.576)}} \right]}{\left[1 + \frac{-9.147 - 4.173 \cdot a_{\max} + 0.652 \cdot M_w}{10.567 + 0.089 \cdot e^{0.089 \cdot (-7.760 \cdot a_{\max} + 78.576)}} \right]} \pm \sigma_{\varepsilon r_d}$$

$$\mathbf{d} < 40 \text{ ft}$$

$$\sigma_{\varepsilon r_d}(d) = d^{0.864} \cdot 0.00814 \quad (4.12)$$

$$\mathbf{d} \geq 40 \text{ ft}$$

$$\sigma_{\varepsilon r_d}(d) = 40^{0.864} \cdot 0.00814 \quad (4.13)$$

4.4.5 Moment Magnitude

Moment magnitude is a value that is usually reported by seismological laboratories following an event and iterated on for a week or two until the final value is set in stone. Calculating the moment magnitude involves an inverse problem to determine the seismic moment. The uncertainty in these calculations comes from the non-unique inversion based on seismograms that are recorded at various teleseismic stations. The dimensions of the fault plane and the amount of slip associated with larger magnitude events tend to be easier to define than with smaller magnitude events. A simple equation Equation 4.14, based on the variance of a series of previous events (1989 Loma Prieta, 1994 Northridge, 1999 Tehuacan, 1999 Kocaeli, 1999 Taiwan, 2001 Denali), was used to estimate this epistemic uncertainty,

$$\sigma_{M_w} \cong 0.5 - 0.45 \cdot \log(M_w) \quad (4.14)$$

4.5 Data Class

After the case histories have been selected and processed they are classified according to the quality of the informational content. Four classes of data are used to group the data, A through D, with D being substandard and therefore not included in the final database. The criteria for the data classes are as follows:

Class A

1. Original CPT trace with q_c and f_s/R_f , using a ASTM D3441 & D5778 spec. cone.
2. No thin layer correction required
3. $\delta_{CSR} \leq 0.20$

Class B

1. Original CPT trace with q_c and f_v/R_f , using a ASTM D3441 & D5778 spec. cone.
2. Thin layer correction.
3. $0.20 < \delta_{CSR} \leq 0.35$

Class C

1. Original CPT trace with q_c and f_v/R_f , but using a non-standard cone (e.g. Chinese cone or mechanical cone).
2. No sleeve data but $FC \leq 5\%$ (i.e. “clean” sand).
3. $0.35 < \delta_{CSR} \leq 0.50$

Class D

1. Not satisfying the criteria for Classes A, B, or C.

4.6 Review Process

The final step in processing the data is an extensive review procedure. Each case in the database is reviewed a minimum of three times. A panel of qualified experts was assembled to do the review, this included in addition to the author and Prof. Raymond B. Seed; Prof. Jon Stewart, Prof. Les Youd, Dr. Rob Kayen, and Prof. Kohji Tokimatsu. Each case was reviewed by the author, Ray Seed, and at least one of the four independent reviewers. The objective was to remove as much human error and epistemic error from the database as possible.

A final note on the review process includes the review of the analytical and statistical procedures. The application of Bayesian analysis to SPT-based liquefaction triggering correlations and the techniques used was reviewed extensively by the Pacific Earthquake Engineering Research Center (PEER), and peer reviewed as journal publications in the Journal of Geotechnical and Geoenvironmental Engineering and the Journal of Structural Reliability. The CPT-based liquefaction triggering correlation, and the associated

Bayesian analysis and methodology, was also reviewed extensively by PEER at quarterly meetings that included as a review panel Prof. Les Youd, Prof. Geoff Martin, and Prof. I.M. Idriss.

It is the author's belief that the power of the Bayesian framework in engineering application is to incorporate all forms of information and that the review process is one of the more important and congenial steps in reducing epistemic uncertainty.

4.7 Conclusion

This chapter includes all the details and procedures used to process data for an unbiased liquefaction triggering correlation within a Bayesian framework. The methods used to generate the best estimates of the representative statistics for each parameter are presented in their entirety. Figures 4.2 through 4.4 show the processed data in $q_{c,1}$ vs. CSR space. The task of developing accurate and appropriate processing techniques was both important and involved, and the final correlation attests to the time well spent.

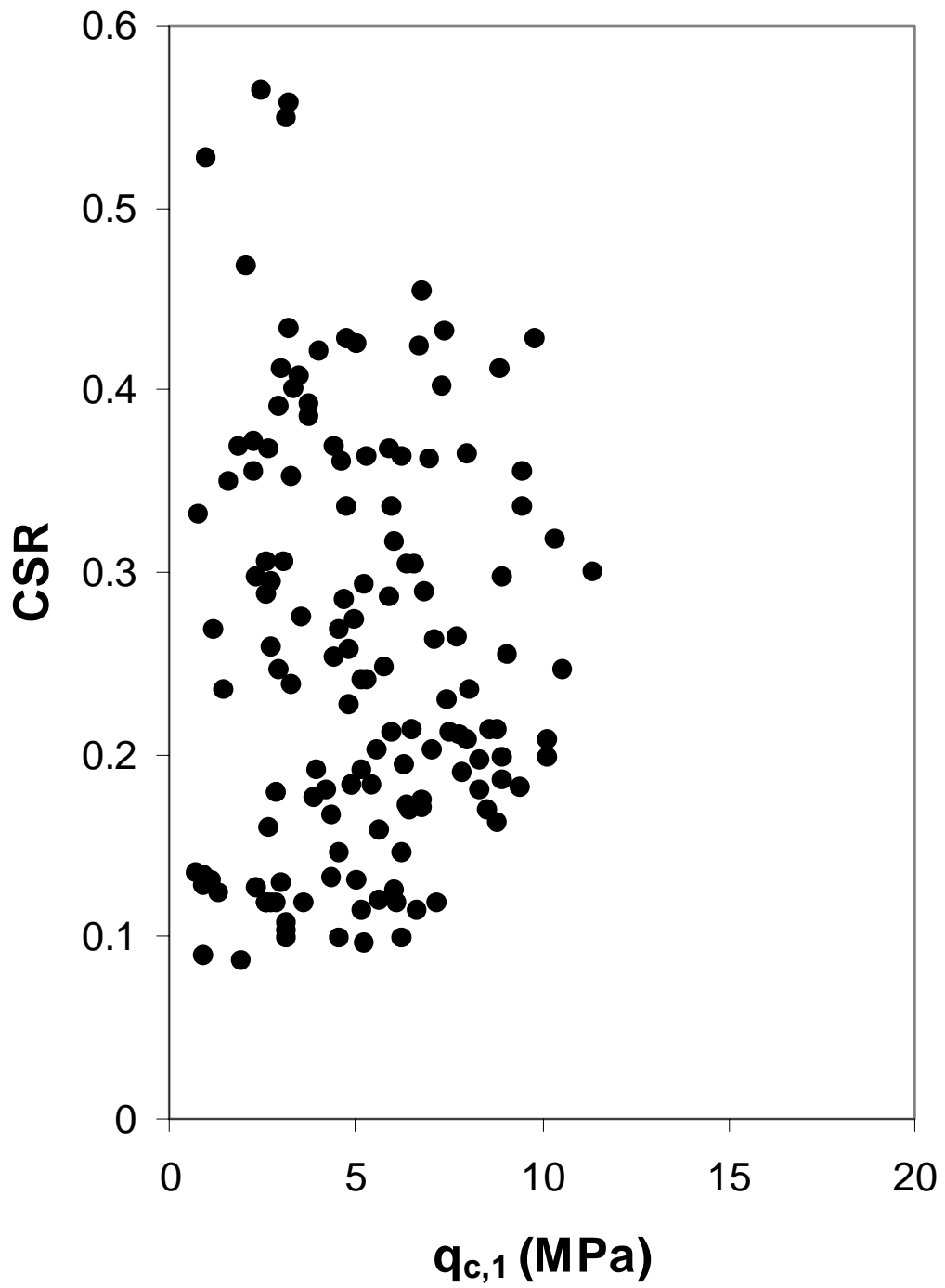


Figure 4.2 Plot Showing Mean Location of Liquefied Data Points.

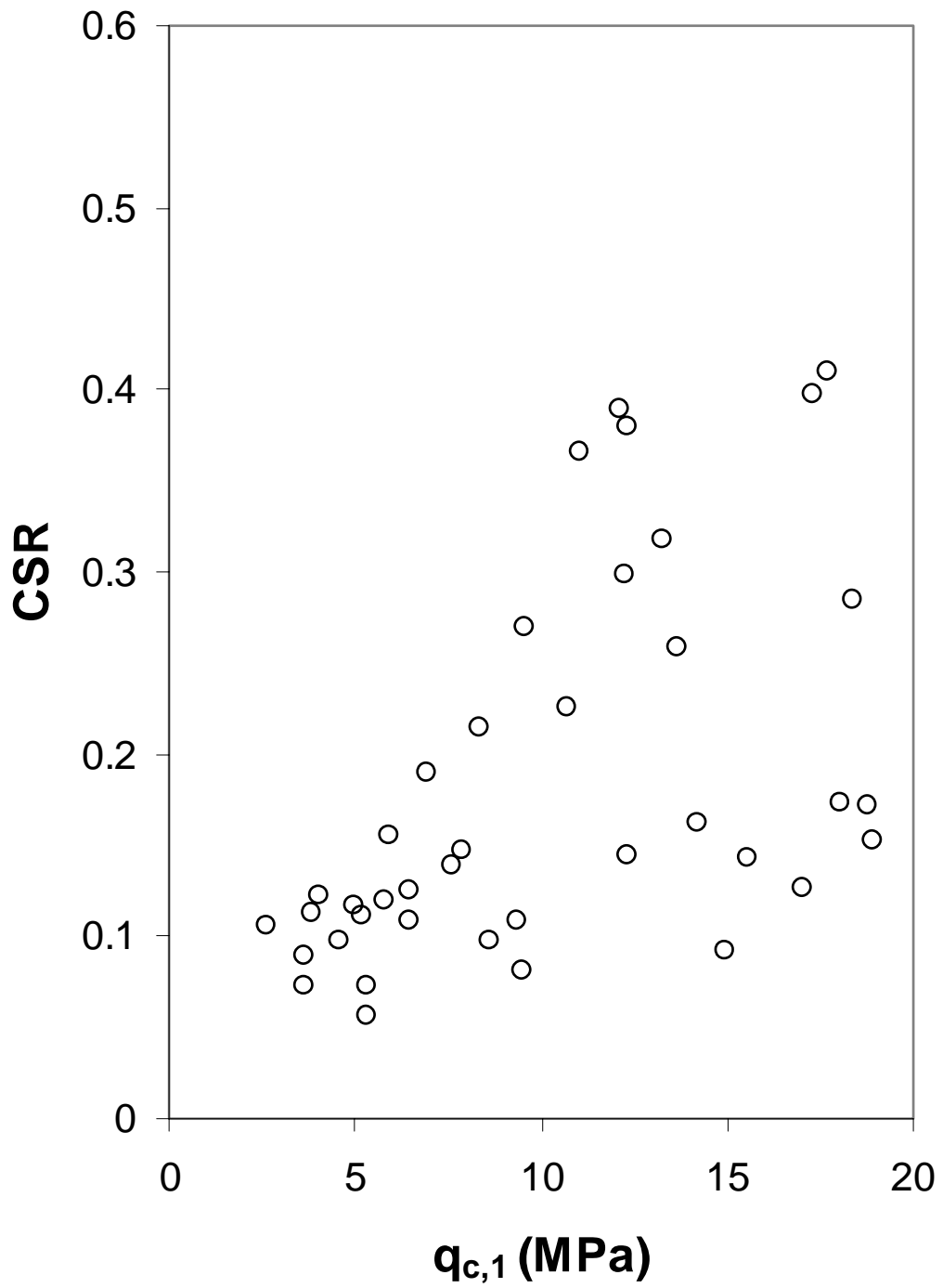


Figure 4.3 Plot Showing Mean Location of Non-Liquefied Data Points.

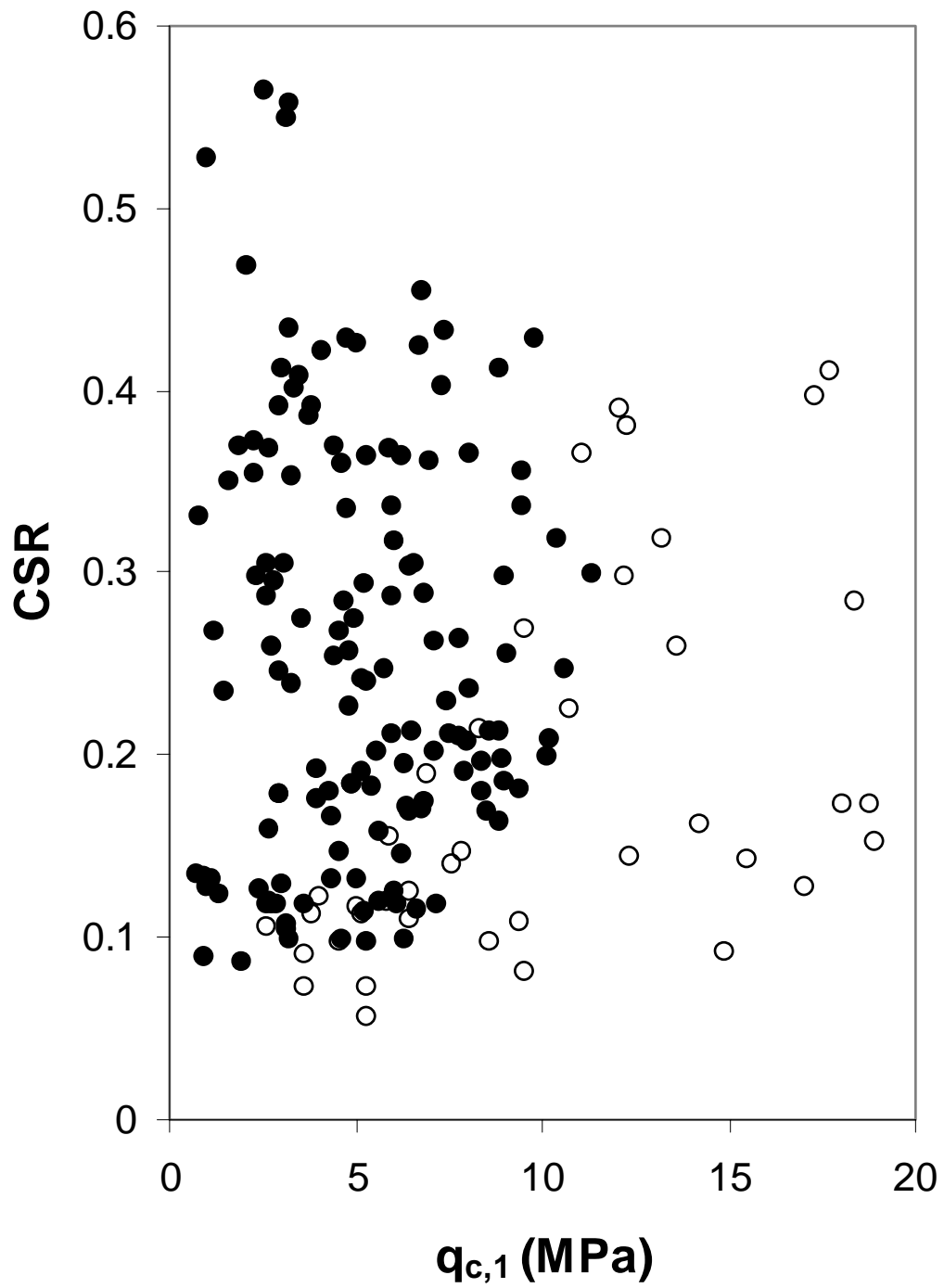


Figure 4.4 Plot Showing Mean Location of Both Liquefied (dots) and Non-Liquefied (circles) Data Points.