Paleoearthquake records typically provide data on large surface-rupturing events which can be difficult to attain via alternative means. These data are the source of key input parameters for seismic hazard models and like many other seismological datasets are subject to sampling limitations and biases. To understand better these sampling issues and to place meaningful interpretations on the available paleoseismic information we have compiled earthquake data for 24 faults each of which ruptured the ground surface in four or more events. Geological data are complemented by earthquakes from numerical simulations and from the historical record over the last ~170 years. Given uncertainties in the data and the relatively small number of observation (24 faults with a total of 158 paleoearthquakes and 20 historical earthquakes of Mw ≥6.5), the results presented are considered preliminary. The magnitude of completeness (Mc) of these events varies with tectonic setting and the rates of surface process. In regions of strike-slip and reverse faulting no events of <Mw 7 ruptured the ground surface historically and Mc values of Mw ~7-7.3 may apply to the paleoseismic data. Given the inferred Mc values the magnitudes are quasi-Characteristic and caution should be exercised when attempting to use these data to differentiate between Characteristic and Gutenberg-Richter earthquake models. In addition, active faults in low strain-rate regions are generally poorly sampled by paleoseismic data in part because fault slip rates are slower than the rates of surface processes and active faults are poorly preserved at the ground surface. In cases where paleoearthquakes ruptured the ground surface (e.g., >Mw ~6-7.3) the recurrence interval (RI) and single-event slip (SES) can vary by more than an order of magnitude on individual faults. Monte Carlo construction of probability density functions for paleoseismic data and numerical simulations suggest that RI values are positively skewed with long recurrence tails (~3 times the mean), while SES are more normally distributed with less variability (coefficient of variation, CV, of 0.6±0.2 versus 0.4±0.2). Variations in RI described by CV values appear to be larger on slower moving faults and when averaged over five or more events are approximately equal to the long-term (10s to 100s kyr) means. Armed with these paleoseismic data and an understanding of their completeness they may be used in seismic hazard analysis for modelling the variability of RI and SES.