Judging Hypocentre Accuracy without Ground Truth

R J Willemann

Many network operators and some data centres that compile data from several networks are implementing new algorithms and travel time models to compute more accurate earthquake hypocentres and hypocentral uncertainties. Greater accuracy is required for rapid response to disasters, research in earth structure, and preparation for on site inspections related to the comprehensive test ban treaty. While the best way to evaluate different hypocentres is comparison with "ground truth", in many cases ground truth locations are unreliable or unknown. An alternative is to compare the residual differences between predicted and observed arrival times for each of the alternative. Difficulty arises, however, in evaluating the statistical significance of residual statistics since their expected distribution is not known and is likely to include outliers. Where two sets of hypocentres are computed using the same method but different travel time models, they will share a common figure of merit (FOM). We show how the two complete sets of FOM values can be compared, so that the statistical significance of their difference can be evaluated without making assumptions about their distribution. When two sets of hypocentres are computed using different location algorithms, they will generally not share a common FOM. For this case, we discuss the utility of robust measures of scale for comparing the residuals directly, rather than the FOM values for each algorithm.