Prediction of quantiles at extreme tails is of interest in numerous applications. Extreme value modelling provides various competing predictors for this point prediction problem. A common method of assessment of a set of competing predictors is to evaluate their predictive performance in a given situation. However, due to the extreme nature of this inference problem, it can be possible that the predicted quantiles are not seen in the historical records, particularly when the sample size is small. This situation poses a problem to the validation of the prediction with its realisation. In this article, we propose two non-parametric scoring approaches to assess extreme quantile prediction mechanisms. The proposed assessment methods are based on predicting a sequence of equally extremal quantiles on different parts of the data. We then use the quantile scoring function to evaluate the competing predictors. The performance of the scoring methods is compared with the conventional scoring method and superiority of the former methods are demonstrated in a simulation study. The methods are then applied to reanalyse cyber Netflow data from Los Alamos National Laboratory and daily precipitation data at a station in California available from Global Historical Climatology Network.