Supplemental Research Data

The following pages contain scatter-plots of estimated versus real accuracy for the Balibase, Prefab, SABmark superfamily and SABmark twilight benchmark. We estimated the accuracy using both MOS and norMD score. The Pearson correlation coefficient (r) is given for each scatter-plot.
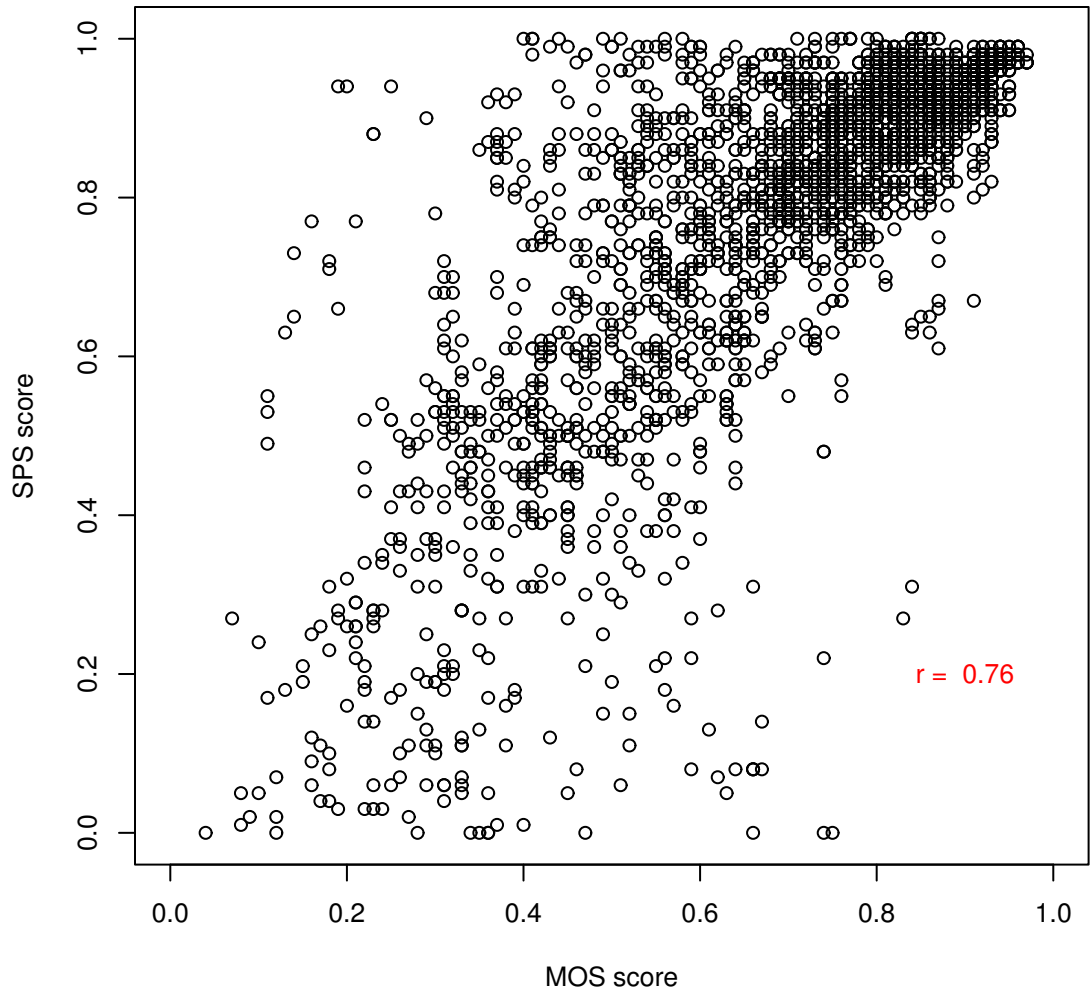
Figure 1: Scatter-plot of estimated accuracy (MOS) versus real accuracy (SPS) on the Balibase test set.
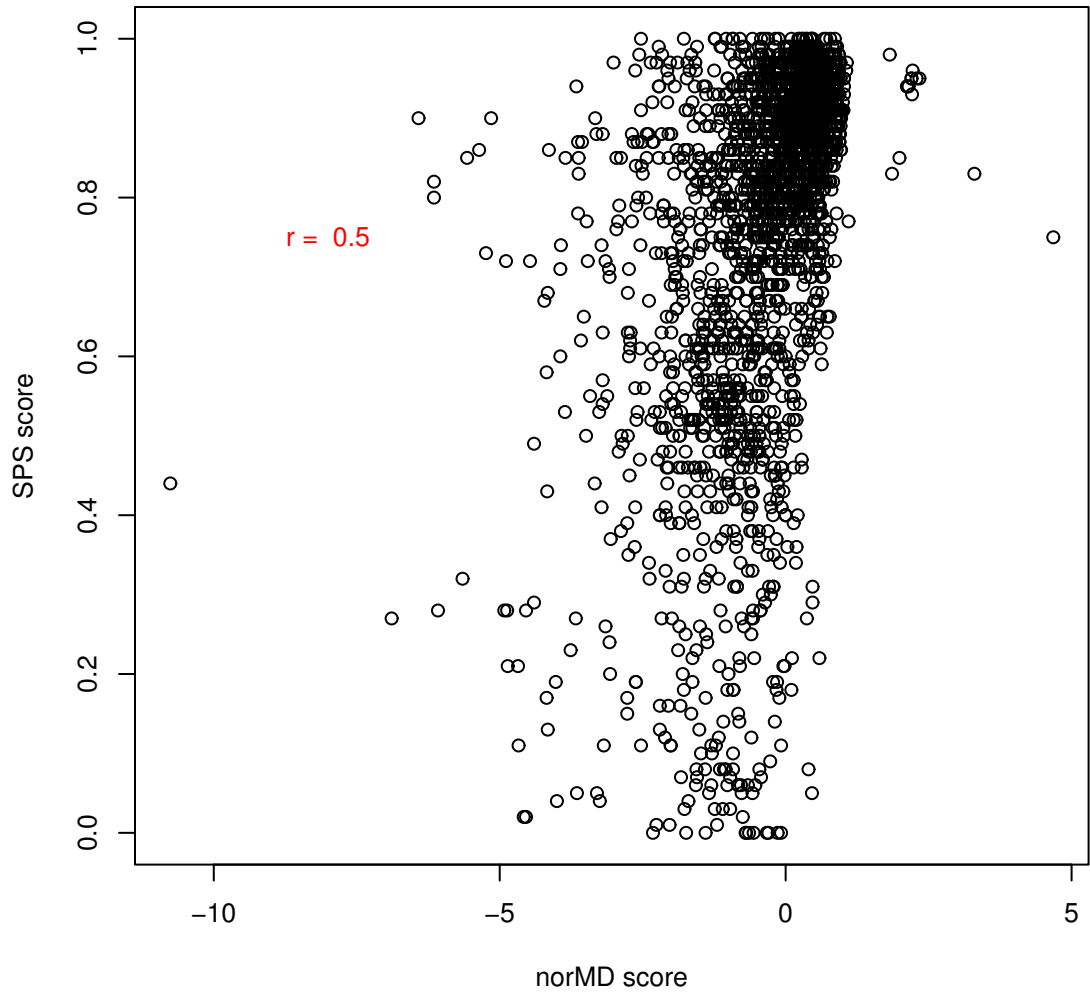
Figure 2: Scatter-plot of estimated accuracy (norMD) versus real accuracy (SPS) on the Balibase test set.
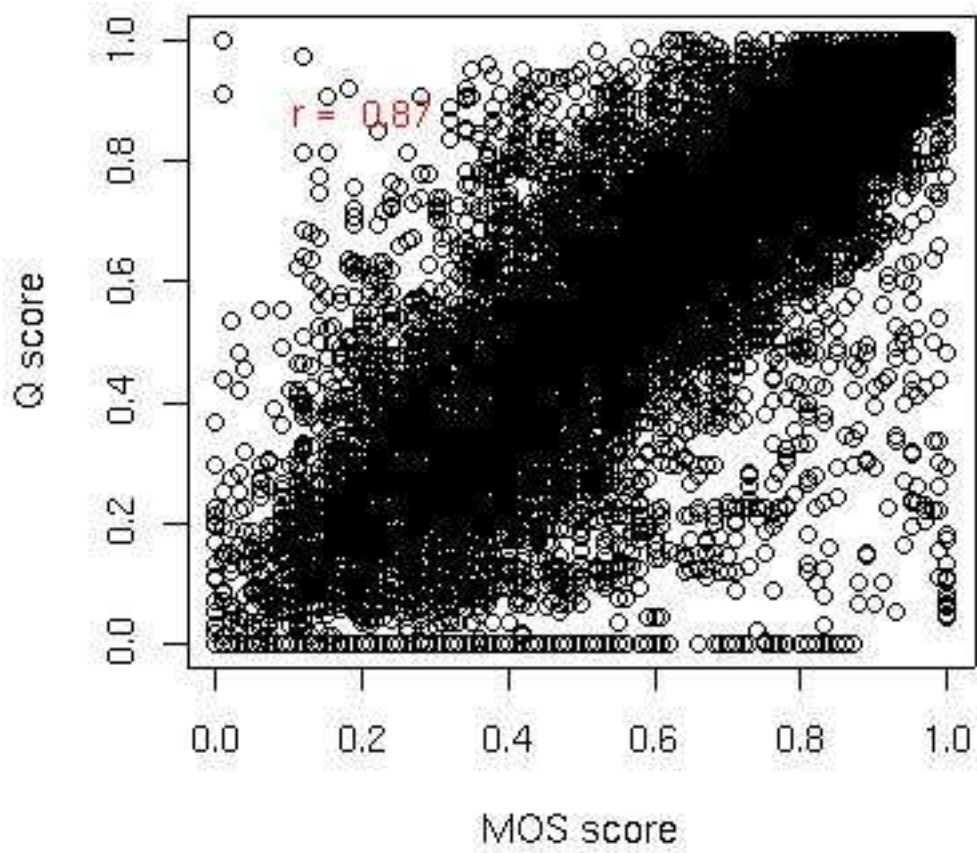
Figure 3: Scatter-plot of estimated accuracy (MOS) versus real accuracy (Q) on the Prefab test set.
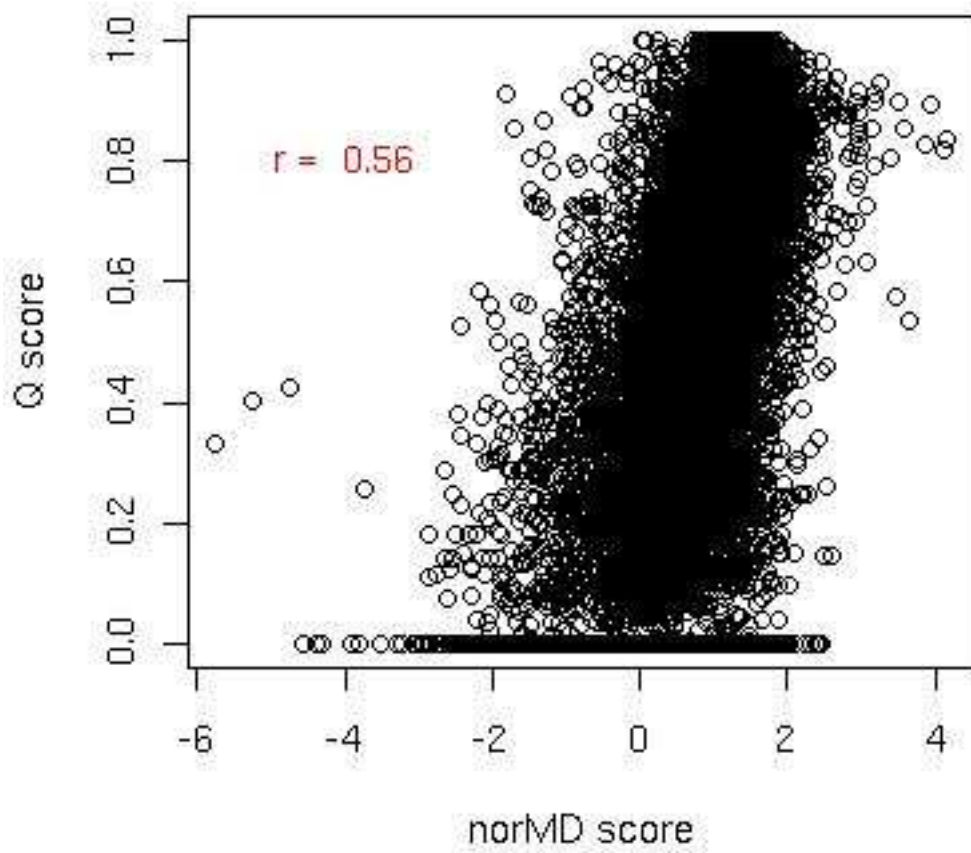
Figure 4: Scatter-plot of estimated accuracy (norMD) versus real accuracy (Q) on the Prefab test set.
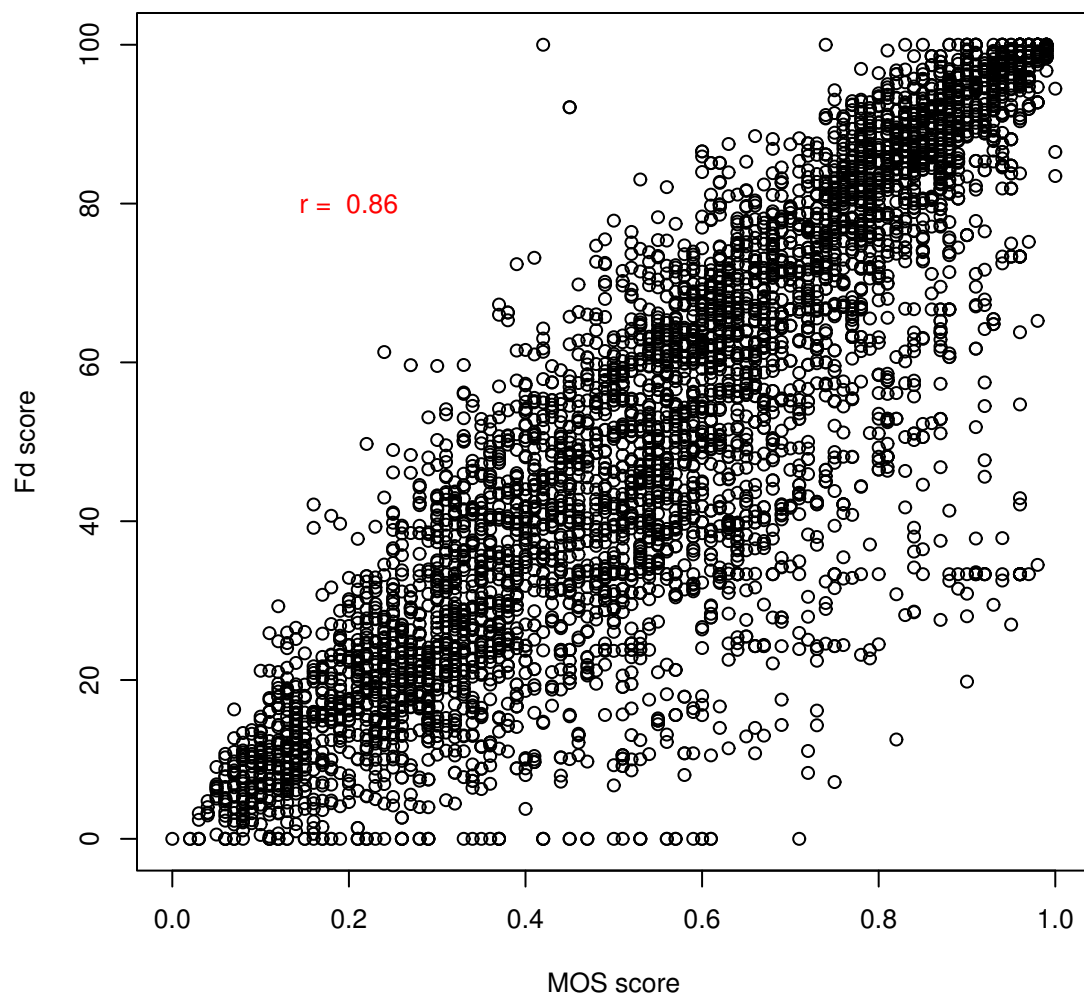
Figure 5: Scatter-plot of estimated accuracy (MOS) versus real accuracy (Fd) on the SABmark superfamily test set.
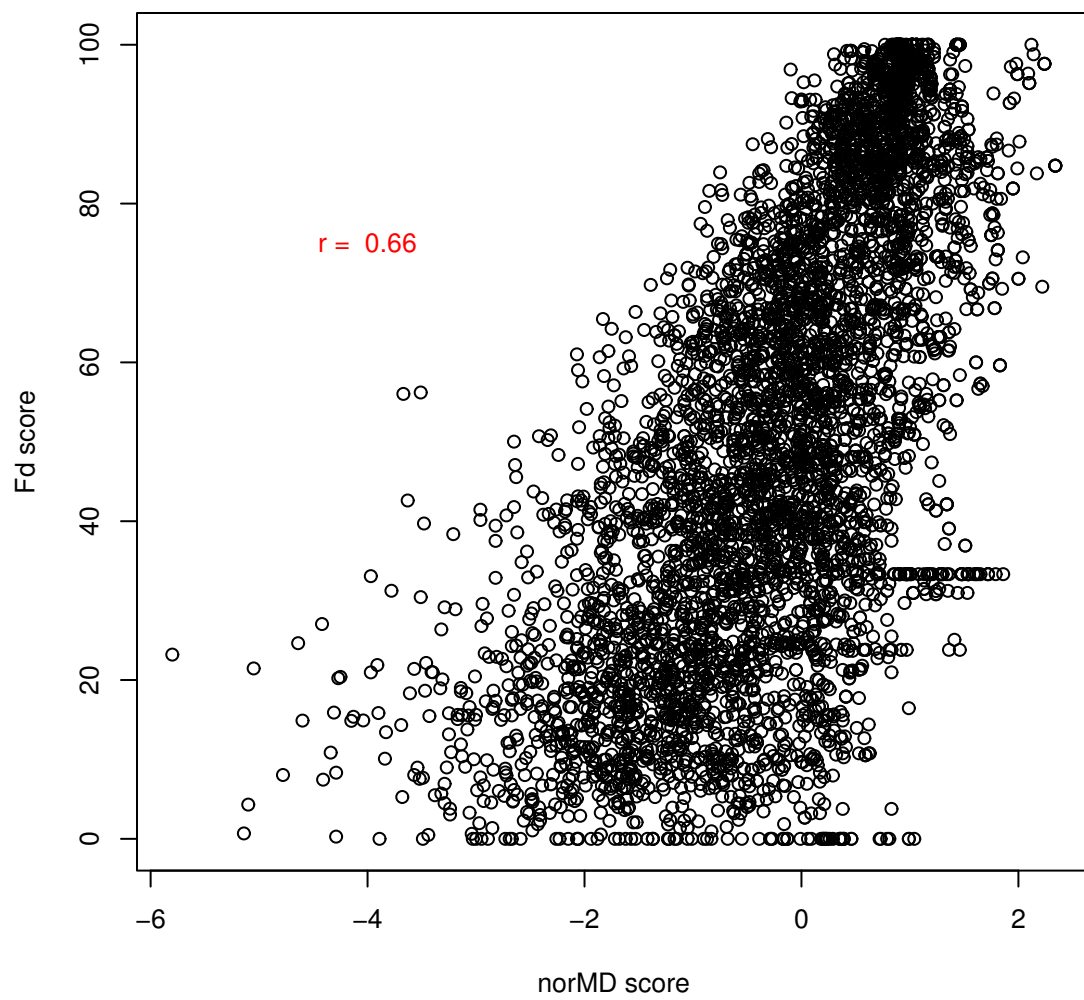
Figure 6: Scatter-plot of estimated accuracy (norMD) versus real accuracy (Fd) on the SABmark superfamily test set.
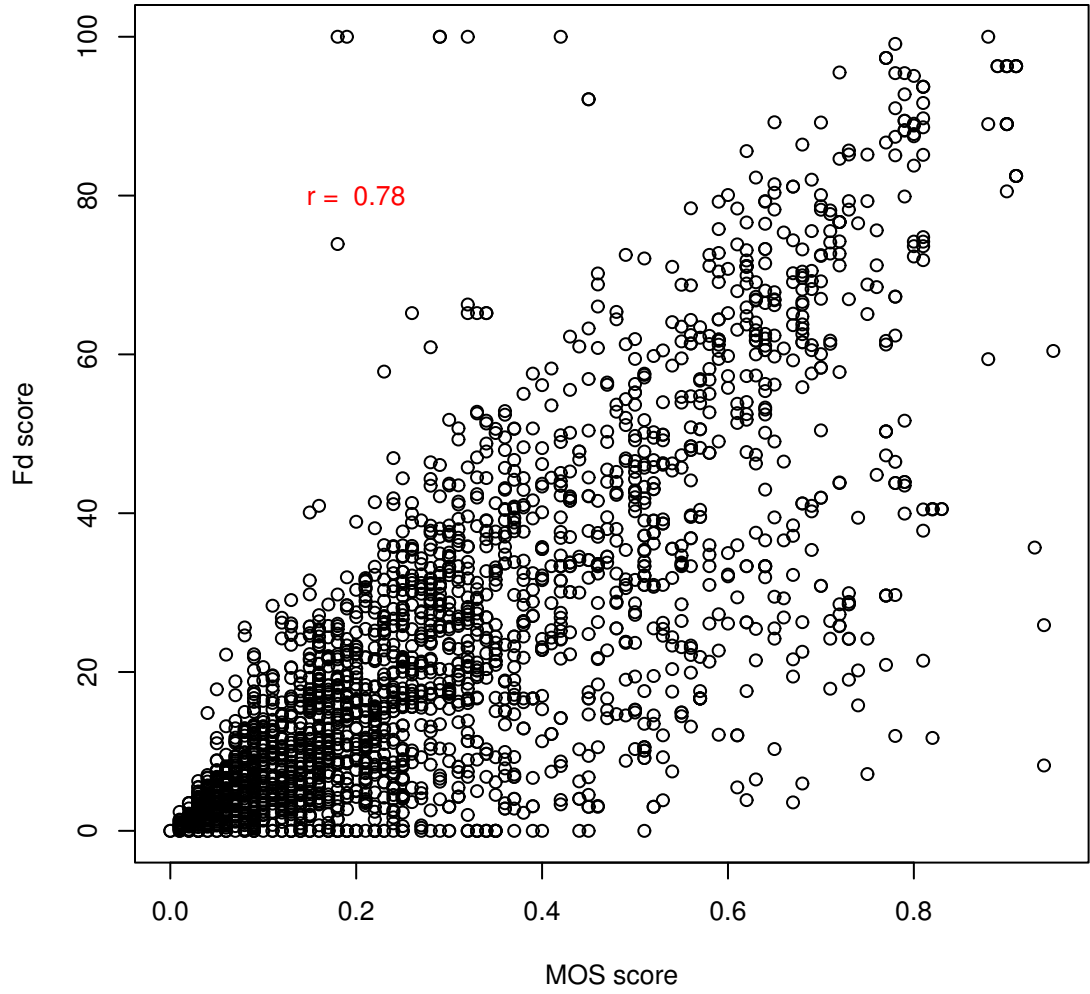
Figure 7: Scatter-plot of estimated accuracy (MOS) versus real accuracy (Fd) on the SABmark twilight test set.
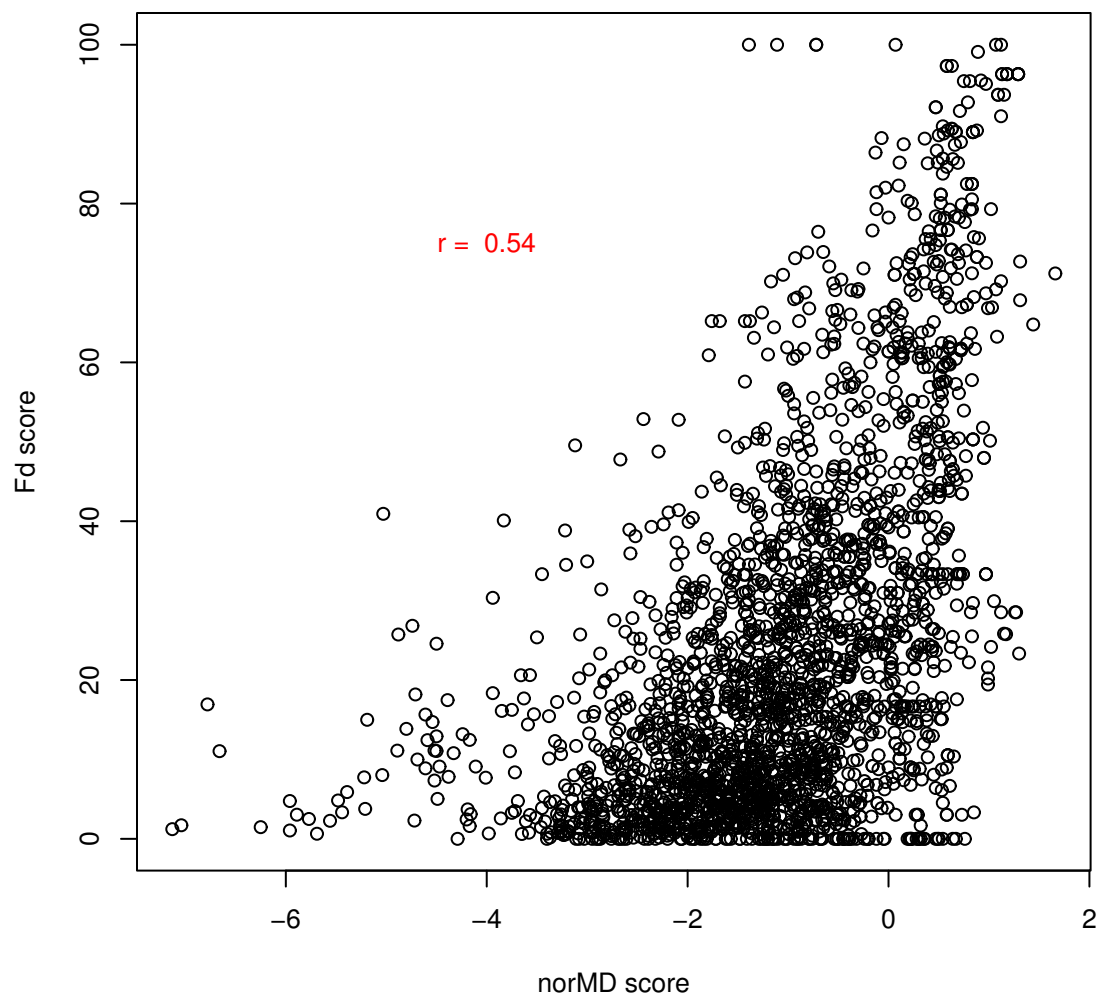
Figure 8: Scatter-plot of estimated accuracy (norMD) versus real accuracy (Fd) on the SABmark twilight test set.