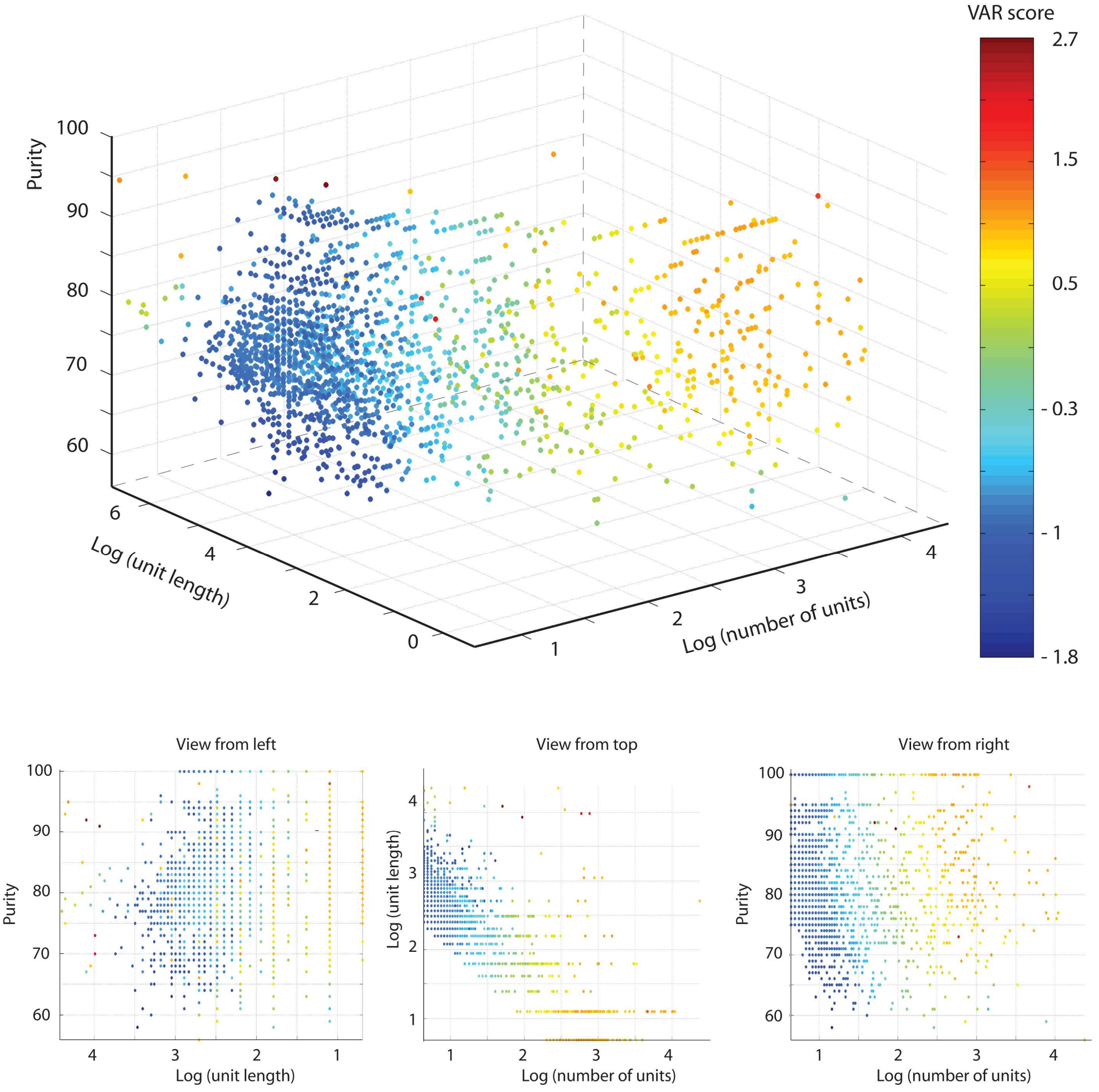

Supplemental Fig. 1. Visualization of the SERV model.

The figure shows the 2,743 tandem repeats used to develop the model (training and validation sets). Each repeat is represented in three dimensional space as a dot with coordinates corresponding to its main features (number of repeat units, unit length in nucleotides and repeat purity). Each dot was colored according to its VARscore (see color scale). Higher VARscores indicate higher predicted variability. The three small plots (2 dimensional views) show how the VARscore and thus the predicted variability increases with increasing repeat unit number, purity and unit length.