Automated Methods for Groove Identification in 3D Bullet Land Scans

Kiegan Rice

Goal: come up with a description for each of the methods and write them out.

Description of Methods

One thing to note is that Methods 2-4 all start with the same base, fitting a quadratic robust linear model. So these could probably be combined into one function with different options if we wanted to…

Method 1: Rollapply

function name: `get_grooves_rollapply` or `get_grooves(method = "rollapply")`

Method 2: Robust Linear Model (4*Median Cutoff)

function name: `get_grooves_quadratic` or `get_grooves(method = "quadratic")`

Steps:

1. Fit a robust linear model of order 2 to the averaged profile of the bullet land. This is fit using the default methods of the 'rlm' function.
2. Calculate the absolute residuals of this model fit.
3. Remove all values that are more than 4 times the median absolute residual value.
4. Extract the range of remaining values. These are the groove locations that are reported.

Method 3: Positive Residuals Cutoff

function name: `get_grooves_quadratic_pos_resid` (this name is gross)

Steps:

1. Fit a robust linear model of order 2 to the averaged profile of the bullet land. This is fit using the default methods of the 'rlm' function.
2. Calculate the residuals (*note: not the absolute residuals*) of this model fit.
3. Remove all points that have positive residuals.
    - Rationale: We expect the grooves to have very high positive residual values, and the robust linear model fit should lead to both positive and negative residuals within the land itself (in the area that we want to focus on).
4. Extract the range of remaining values. These are the groove locations that are reported.

Method 4: Median Absolute Deviation Cutoff

function name: `get_grooves_mad` (not consistent - should probably have a quadratic somewhere in there)
Steps:

1. Fit a robust linear model of order 2 to the averaged profile of the bullet land. This is fit using the default methods of the ‘rlm‘ function.
2. Calculate the absolute residuals of this model fit.
3. Calculate the median absolute deviation of the absolute residuals (using the ‘mad‘ function).
4. Remove all points that have an absolute residual higher than the median absolute deviation.
   - Rationale: The median absolute deviation might give us a good cutoff because it would remove high outlying residual values (like those we expect to see in the grooves)
   - Question: Do we actually want to use the MAD itself?
5. Extract the range of remaining values. These are the groove locations that are reported.

Method 5: Predicting Groove Sides

function name: `get_grooves_iterate2`

Steps:

1. Fit a robust linear model of order 2 to the averaged profile of a bullet land.
2. Sum residuals on left side - from min to 600
   - Rationale: The furthest in groove on the left side for Hamby44 is at 543.
3. Sum residuals on right side - from 1800 to max
   - Rationale: The furthest in groove on the right side for Hamby44 is at 1872
4. Iteratively remove data from the left side (1% at a time) until the sum of the residuals on the left side is less than or equal to 200.
5. Iteratively remove data from the right side (1% at a time) until the sum of the residuals on the right side is less than or equal to 100.

Drawbacks:

1. Currently, extremely dependent on robust linear model fit - all of our methods are.
2. Cuts off a lot of data if the robust linear model fit isn't good - there is a high sum of residuals.

Method 6: Robust Loess Fit (2*Median Cutoff)

function name: `get_grooves_robust_loess_cutoff`

Steps:

1. Fit a robust loess model (span = 1) to the averaged profile of a bullet land.
2. Calculate the 2*median absolute residual from the robust loess model.
3. Remove all absolute residuals greater than 2*median absolute residual.
4. Find the range of the remaining Y values - these are the grooves!

Metrics for Comparing Performance of Methods

Metric 1: Sums of Absolute Residuals Between True Groove, Predicted Groove - Robust Linear Model Fit

Steps:

1. Identify predicted groove on each side.
2. Store the residuals from the robust linear model.
3. Extract the range of y values *between* the predicted left groove and manually identified left groove.
4. Sum the residuals that occur within that range of y values.
5. Repeat steps 3-4 for the right groove.
6. Two methods of looking at these:
   (a) Repeat for different methods, and compare the sums across the same set of lands.
   (b) Repeat for different methods, and create comparative boxplots (again, across the same set of lands) to visually investigate whether methods are reducing large errors.

Drawbacks:

1. Doesn't treat "underpredicting" (not cutting off groove) and "overpredicting" (cutting off extra bullet land) differently - just sums residuals.
2. Is dependent on the robust linear model, because it uses the residuals from that model. So when that model is a bad fit, we will get lots of wonky things anyways.

Metric 2: Sums of Absolute Residuals Between True Groove, Predicted Groove - Robust Loess Model Fit

Steps:

1. Identify predicted groove on each side.
2. Store the residuals from the robust loess model.
3. Extract the range of y values *between* the predicted left groove and manually identified left groove.
4. Sum the residuals that occur within that range of y values.
5. Repeat steps 3-4 for the right groove.
6. Two methods of looking at these:
   (a) Repeat for different methods, and compare the sums across the same set of lands.
   (b) Repeat for different methods, and create comparative boxplots (again, across the same set of lands) to visually investigate whether methods are reducing large errors.

Drawbacks:

1. Same drawbacks as Metric 1, but could potentially be a better initial fit than the robust linear model.

Metric 3: Distance from the predicted groove to the manually identified groove