

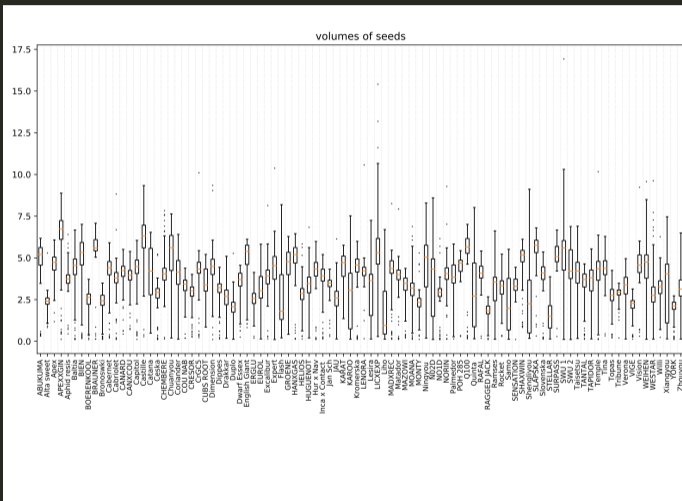
# Looking at microCT data of Brassica pods

I am not a biologist, please stop me and correct me if I say silly things.

# Pod Width

# Sphericity

# Volume



# Surface Area

# Correlations

# Filtering false seeds



- Image analysis produces many false seeds at the beak tip
- Density and size is comparable to seed
- Hard to recognise by graphical methods alone
- Recognise them by mathematical means instead

# Spine fitting

- For every CT slice we have the centroid of the object
- Fit X and Y position as cubic functions of z
- Define 'real z' as the distance measured along the fitted curve from the beak to the z coordinate of the point





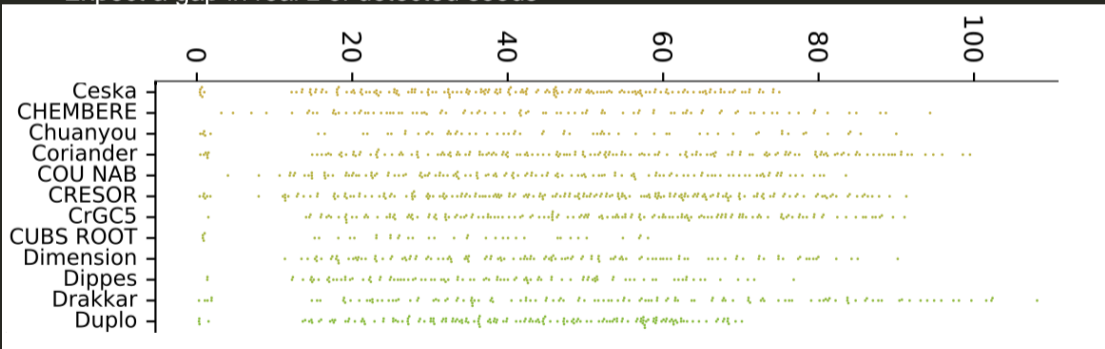
# Distinguishing between beak tip and Real Seeds™

## Failed approaches:

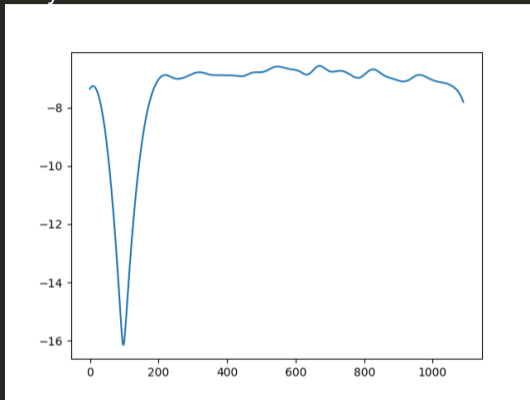
1. Assert that seeds must not be implausible - Removed insufficiently many seeds
  - Too close to the ends of the pod
  - Too large given pod dimensions
2. Real z position of seeds of a pod is a sample from some probability distribution, fit and parameterize the distribution to classify seeds.
  - Sum of two normal(-ish) distributions - noise at beak might be normal, everything else definitely is not
  - More complicated distribution - too complicated
3. K-Means clustering - Silly for 1 dimensional data
4. Jenks Natural Breaks Optimisation - Should work in theory, did not work well in practice

# Break at Minimum Kernel Density Estimation (KDE)

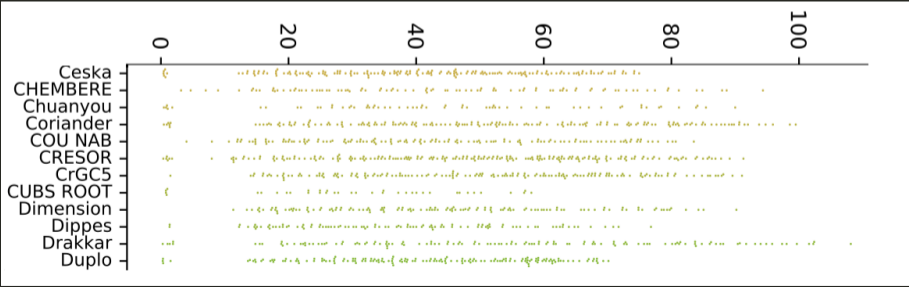
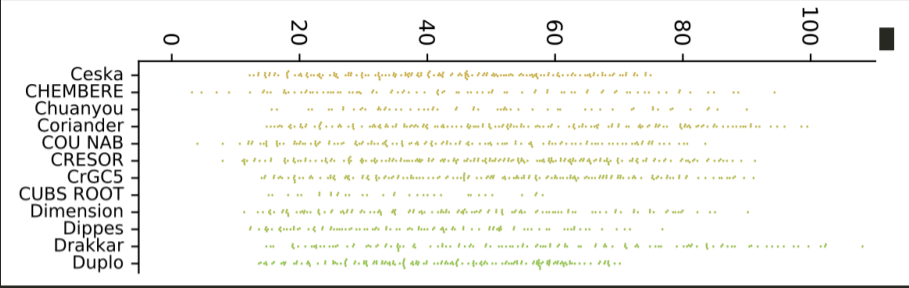
- Beak has no Real Seeds™ and low density
- Expect a gap in real z of detected seeds



- Use KDE to find density of seeds as function of real  $z$



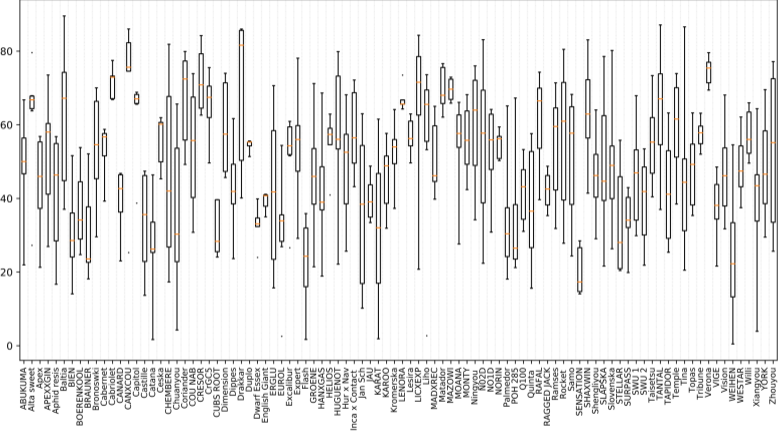
- First seed has real  $z$  less than 100?
- Find the local minimum at lowest real  $z$  where  $\log(\text{KDE}) < -10$
- Keep seeds with greater real  $z$
- Profit



# Beak and Silique length

Use the seed with lowest real  $z$  to mark the boundary of beak and silique:

silique length



beak length

