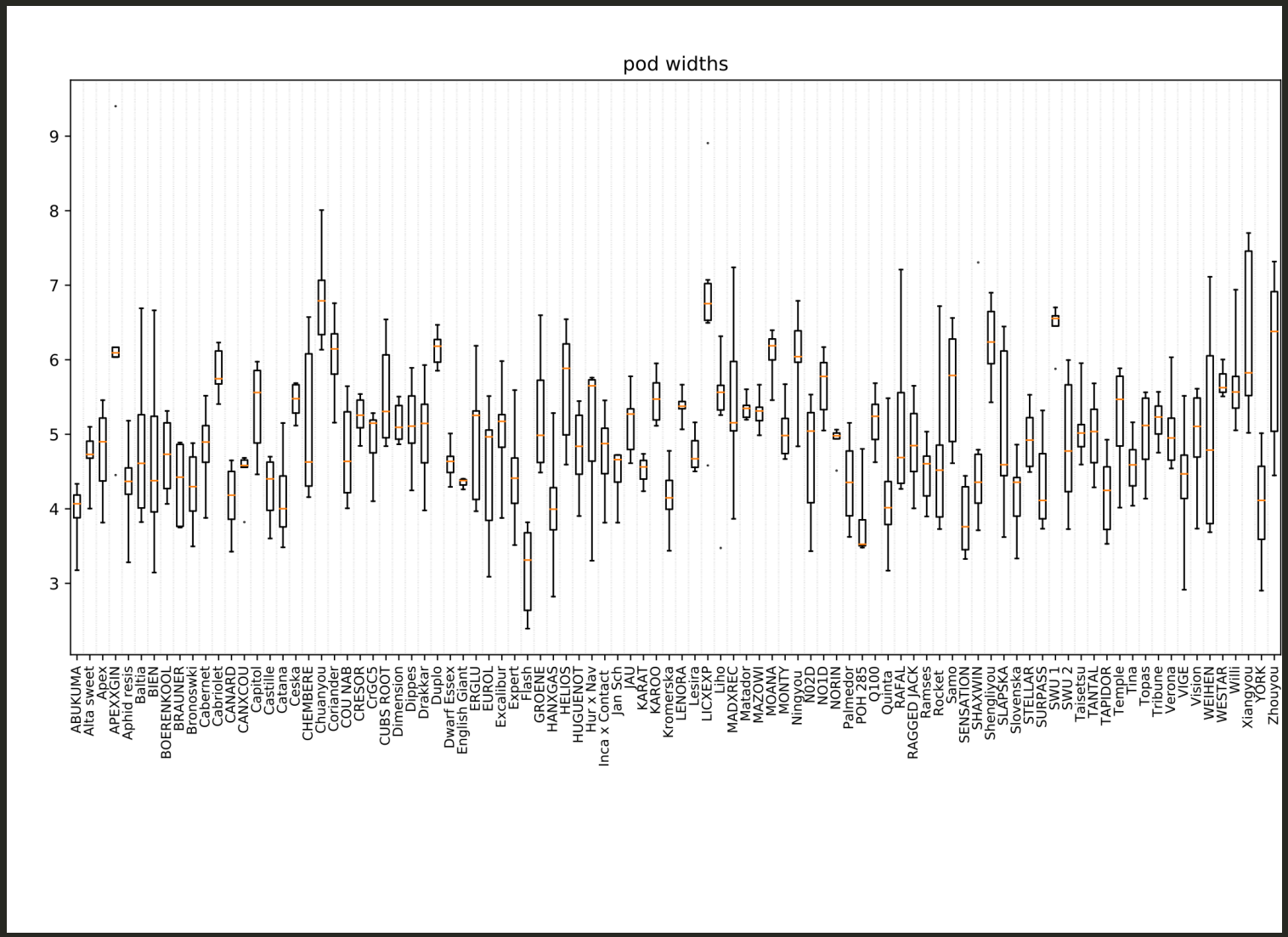
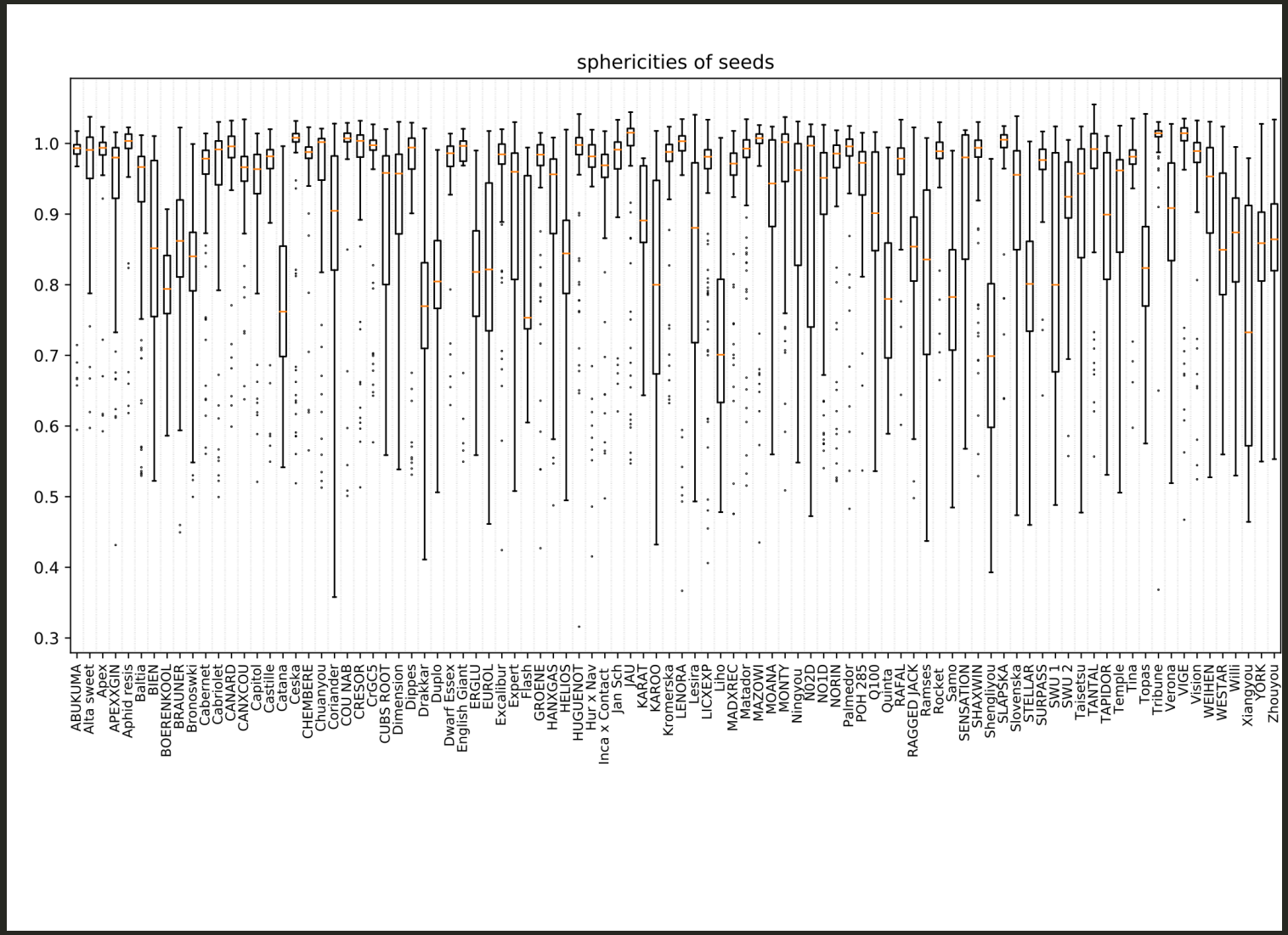
Looking at microCT data of Brassica pods

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

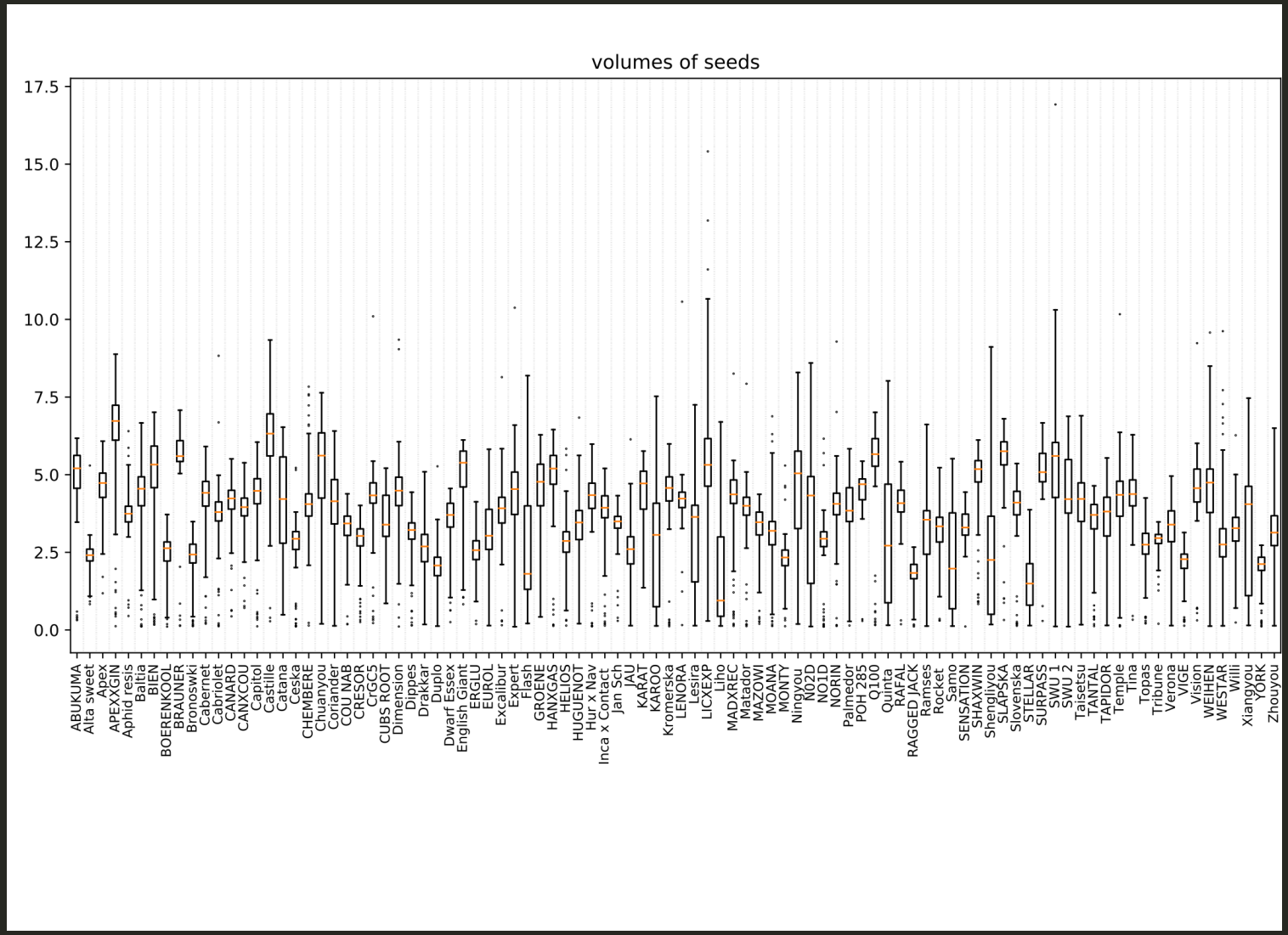
Pod Width


pod widths
9
8
7
6
5
4
3
ABUKUMA
Alta sweet
Apex
APEXXGIN
Aphid resis
Balita
BIEN
BOERENKOOL
BRAUNER
Bronoswki
Cabernet
Cabriolet
CANARD
CANXCOU
Capitol
Castille
Catana
Ceska
CHEMBERE
Chuanyou
Coriander
COU NAB
CRESOR
CrGC5
CUBS ROOT
Dimension
Dippes
Drakkar
Duplo
Dwarf Essex
English Giant
ERGLU
EUROL
Excalibur
Expert
Flash
GROENE
HANXGAS
HELIOS
HUGUENOT
Hur x Nav
Inca x Contact
Jan Sch
JAU
KARAT
KAROO
Kromerska
LENORA
Lesira
LICXEXP
Liho
MADXREC
Matador
MAZOWI
MOANA
MONTY
Ningyou
N02D
NO1D
NORIN
Palmedor
POH 285
Q100
Quinta
RAFAL
RAGGED JACK
Ramses
Rocket
Samo
SENSATION
SHAXWIN
Shengliyou
SLAPSKA
Slovenska
STELLAR
SURPASS
SWU 1
SWU 2
Taisetsu
TANTAL
TAPIDOR
Temple
Tina
Topas
Tribune
Verona
VIGE
Vision
WEIHEN
WESTAR
Willi
Xiangyou
YORK
Zhouyou

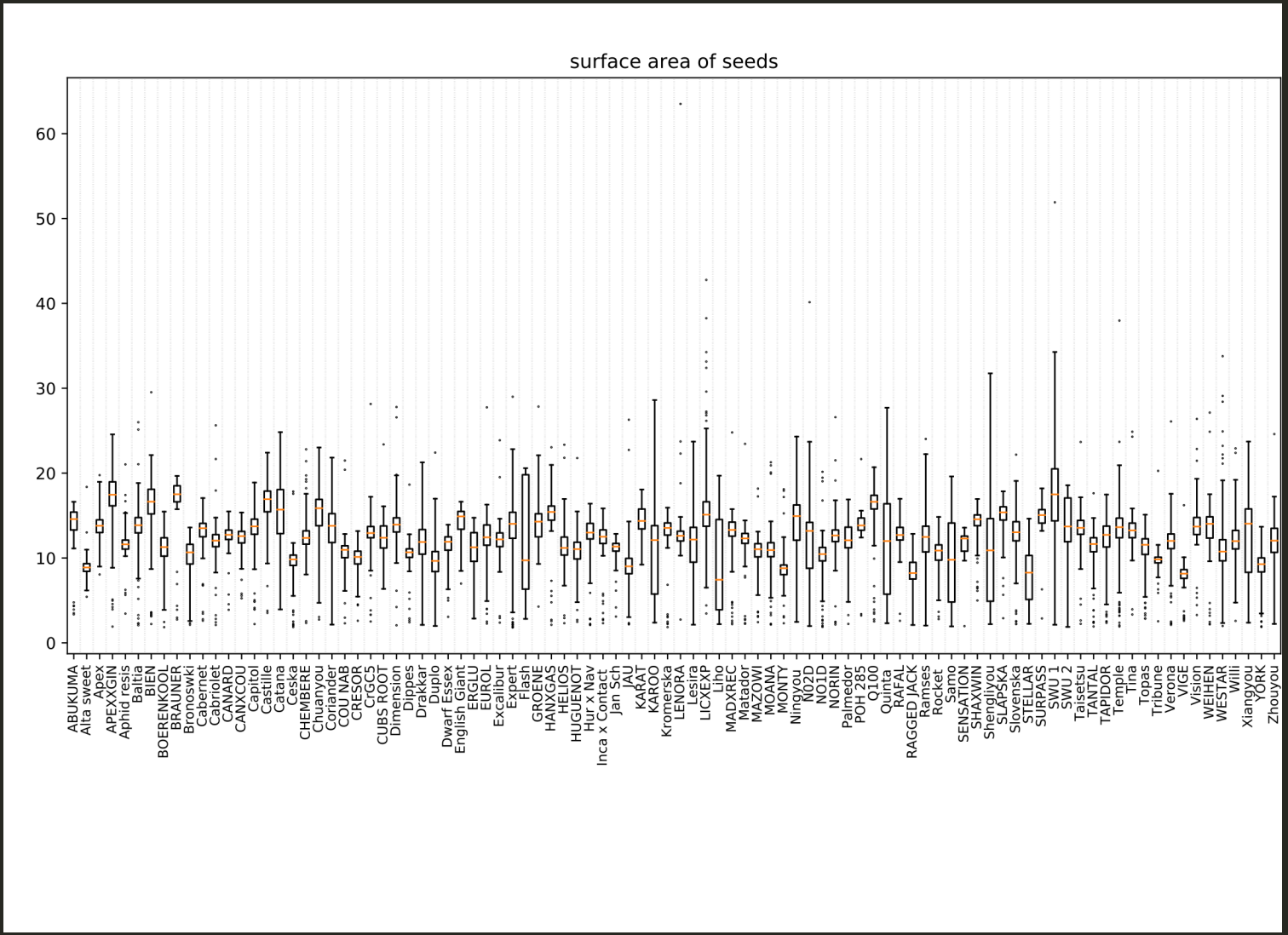
Sphericity


sphericities of seeds
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
ABUKUMA
Alta sweet
Apex
APEXXGIN
Aphid resis
Balita
BIEN
BOERENKOOL
BRAUNER
Bronoswki
Cabernet
Cabriolet
CANARD
CANXCOU
Capitol
Castille
Catana
Ceska
CHEMBERE
Chuanyou
Coriander
COU NAB
CRESOR
CrGC5
CUBS ROOT
Dimension
Dippes
Drakkar
Duplo
Dwarf Essex
English Giant
ERGLU
EUROL
Excalibur
Expert
Flash
GROENE
HANXGAS
HELIOS
HUGUENOT
Hur x Nav
Inca x Contact
Jan Sch
JAU
KARAT
KAROO
Kromerska
LENORA
Lesira
LICXEXP
Liho
MADXREC
Matador
MAZOWI
MOANA
MONTY
Ningyou
N02D
NO1D
NORIN
Palmedor
POH 285
Q100
Quinta
RAFAL
RAGGED JACK
Ramses
Rocket
Samo
SENSATION
SHAXWIN
Shengliyou
SLAPSKA
Slovenska
STELLAR
SURPASS
SWU 1
SWU 2
Taisetsu
TANTAL
TAPIDOR
Temple
Tina
Topas
Tribune
Verona
VIGE
Vision
WEIHEN
WESTAR
Willi
Xiangyou
YORK
Zhouyou

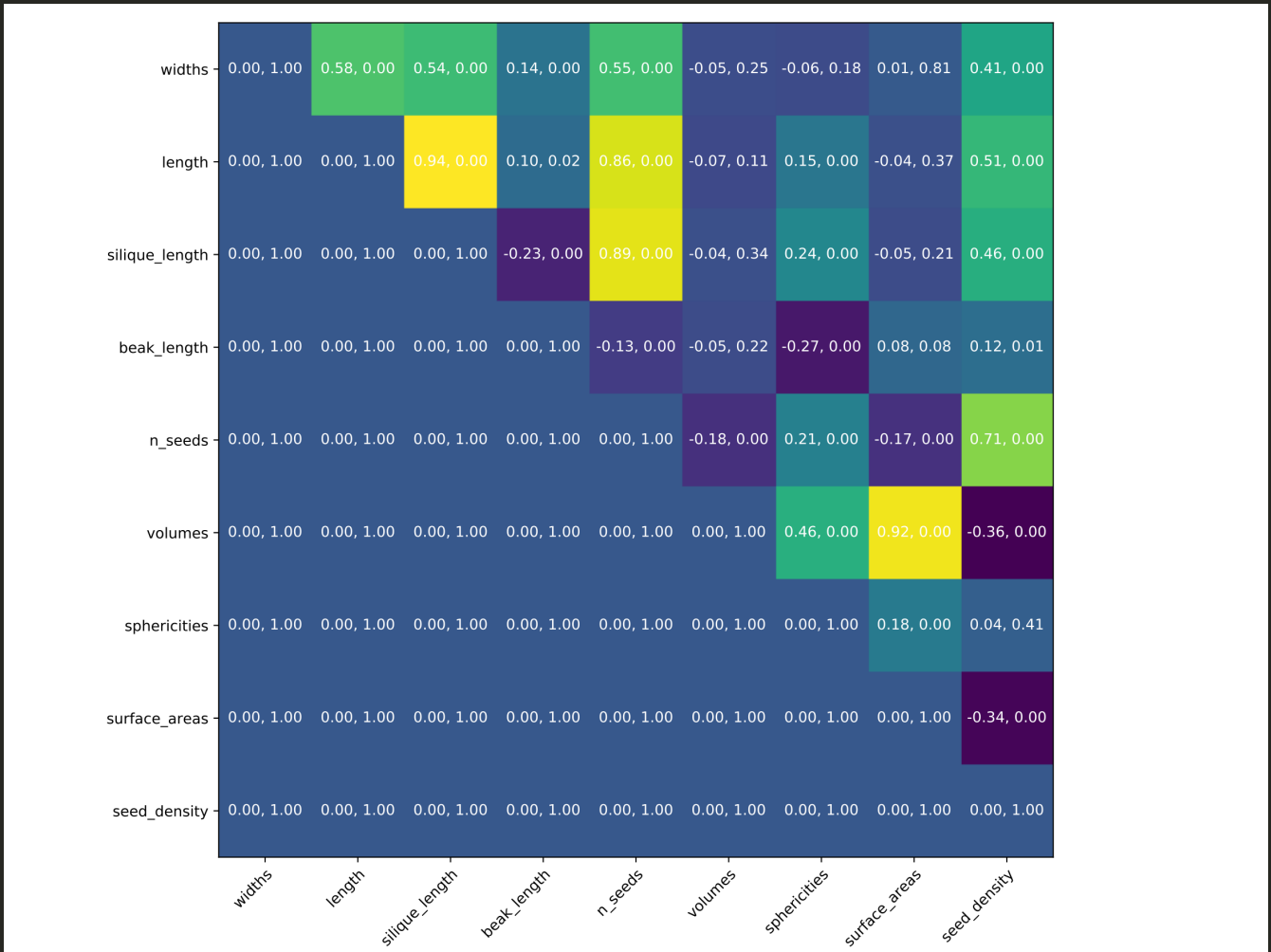
Volume

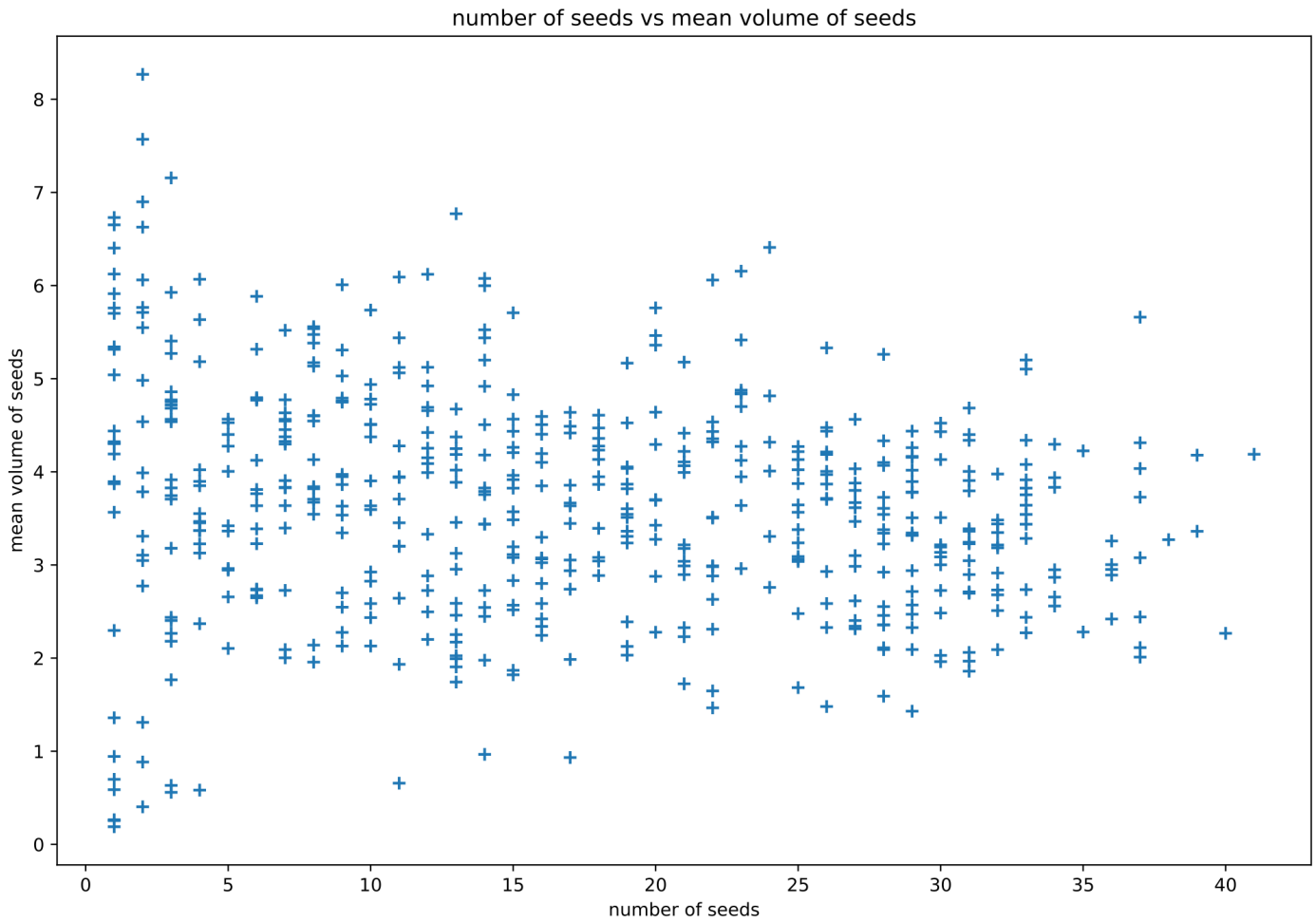

volumes of seeds
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
ABUKUMA
Alta sweet
Apex
APEXXGIN
Aphid resis
Balita
BIEN
BOERENKOOL
BRAUNER
Bronoswki
Cabernet
Cabriolet
CANARD
CANXCOU
Capitol
Castille
Catana
Ceska
CHEMBERE
Chuanyou
Coriander
COU NAB
CRESOR
CrGC5
CUBS ROOT
Dimension
Dippes
Drakkar
Duplo
Dwarf Essex
English Giant
ERGLU
EUROL
Excalibur
Expert
Flash
GROENE
HANXGAS
HELIOS
HUGUENOT
Hur x Nav
Inca x Contact
Jan Sch
JAU
KARAT
KAROO
Kromerska
LENORA
Lesira
LICXEXP
Liho
MADXREC
Matador
MAZOWI
MOANA
MONTY
Ningyou
N02D
NO1D
NORIN
Palmedor
POH 285
Q100
Quinta
RAFAL
RAGGED JACK
Ramses
Rocket
Samo
SENSATION
SHAXWIN
Shengliyou
SLAPSKA
Slovenska
STELLAR
SURPASS
SWU 1
SWU 2
Taisetsu
TANTAL
TAPIDOR
Temple
Tina
Topas
Tribune
Verona
VIGE
Vision
WEIHEN
WESTAR
Willi
Xiangyou
YORK
Zhouyou

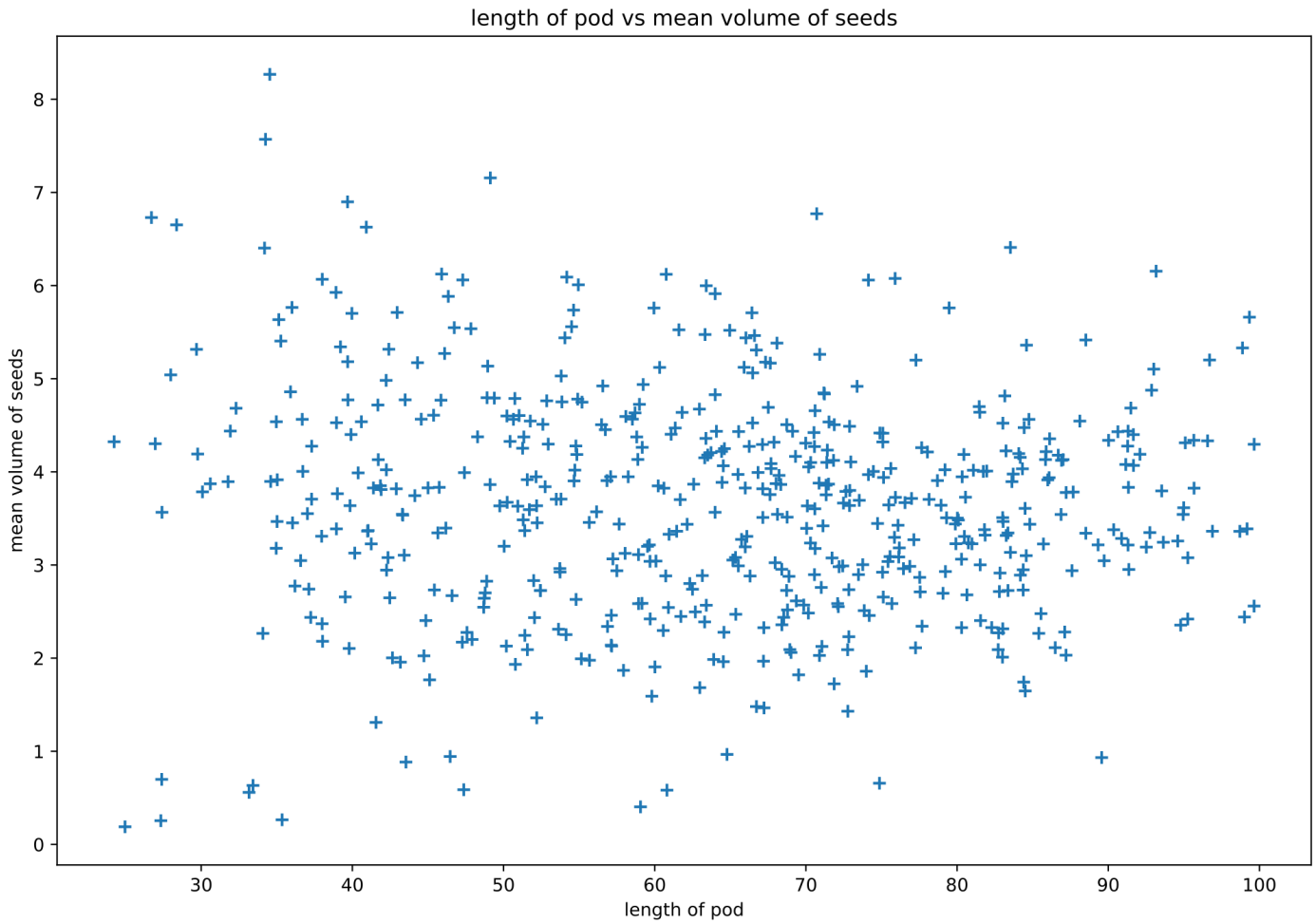
Surface Area

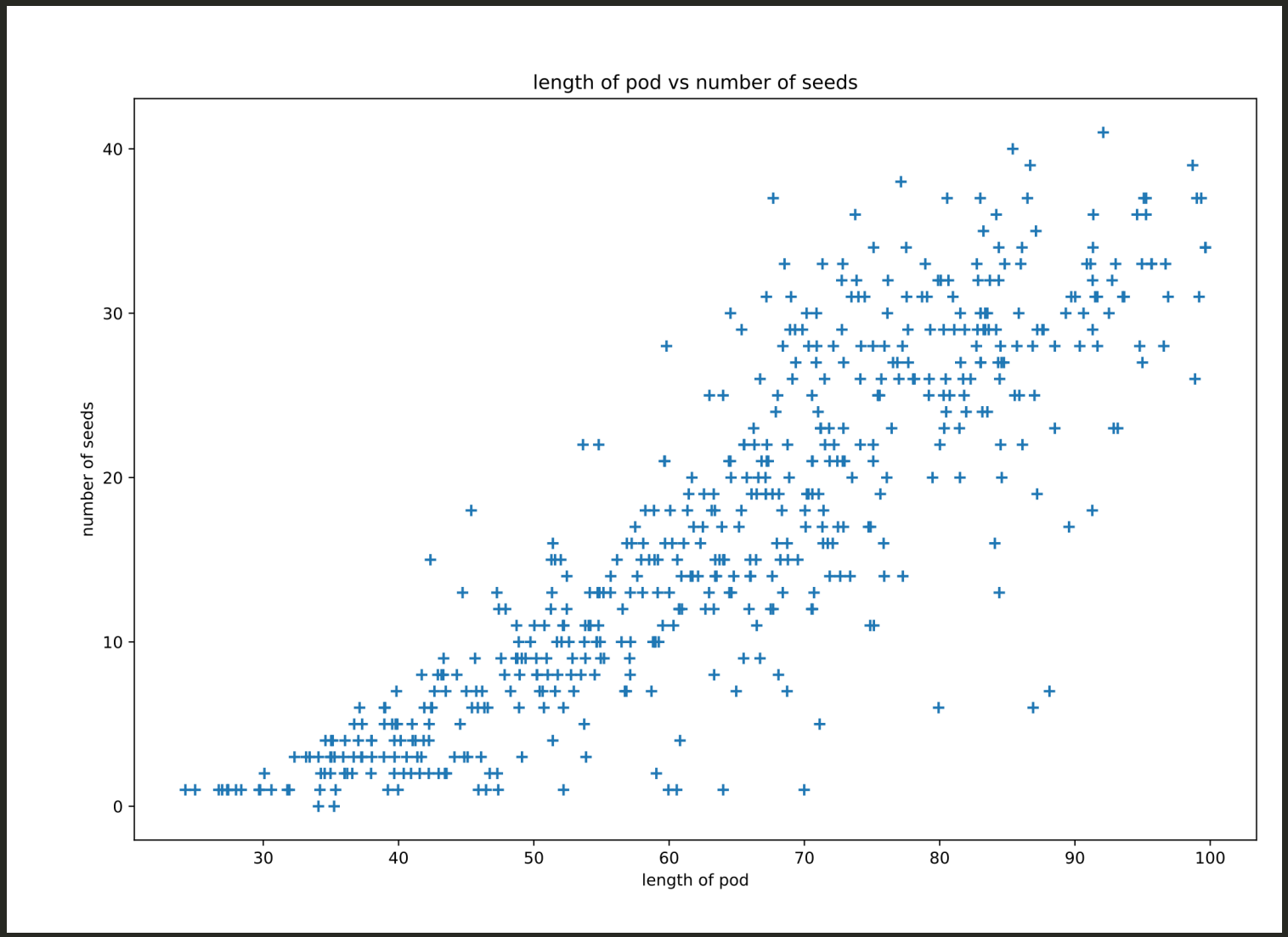


Correlations




number of seeds vs mean volume of seeds
mean volume of seeds
number of seeds
0
1
2
3
4
5
6
7
8
0
5
10
15
20
25
30
35
40


length of pod vs mean volume of seeds
mean volume of seeds
length of pod
0
1
2
3
4
5
6
7
8
30
40
50
60
70
80
90
100

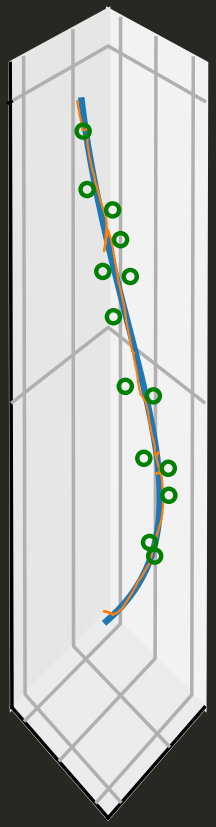

length of pod vs number of seeds
number of seeds
length of pod
0
10
20
30
40
30
40
50
60
70
80
90
100

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



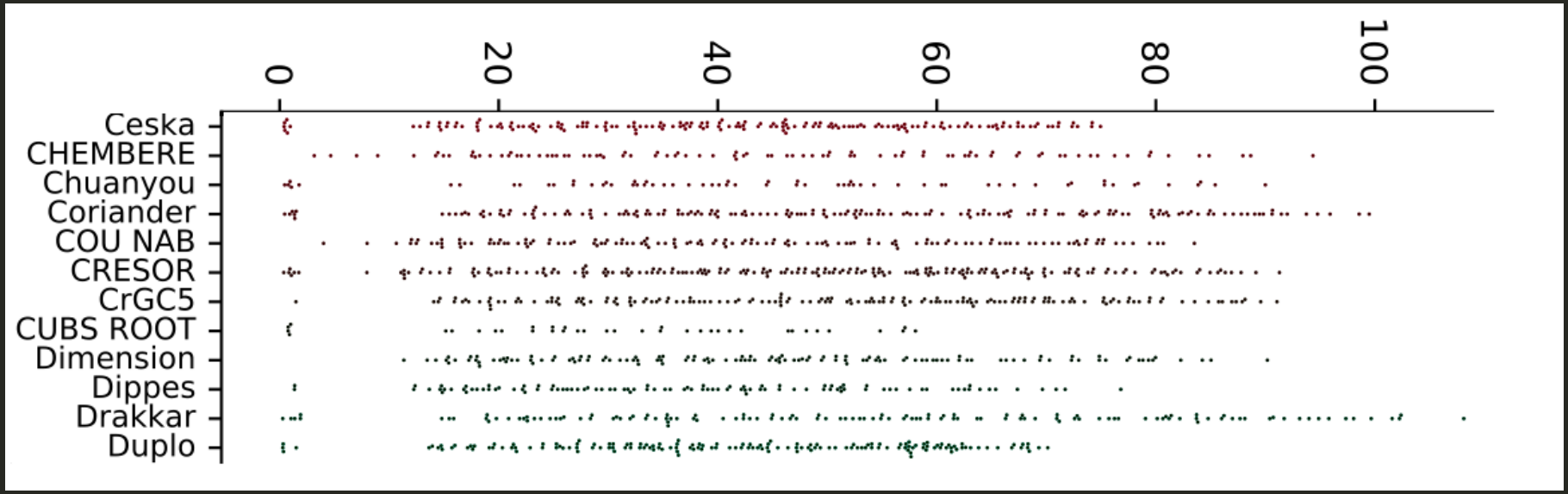
Classify beak tip and Real Seeds™

Failed approaches:

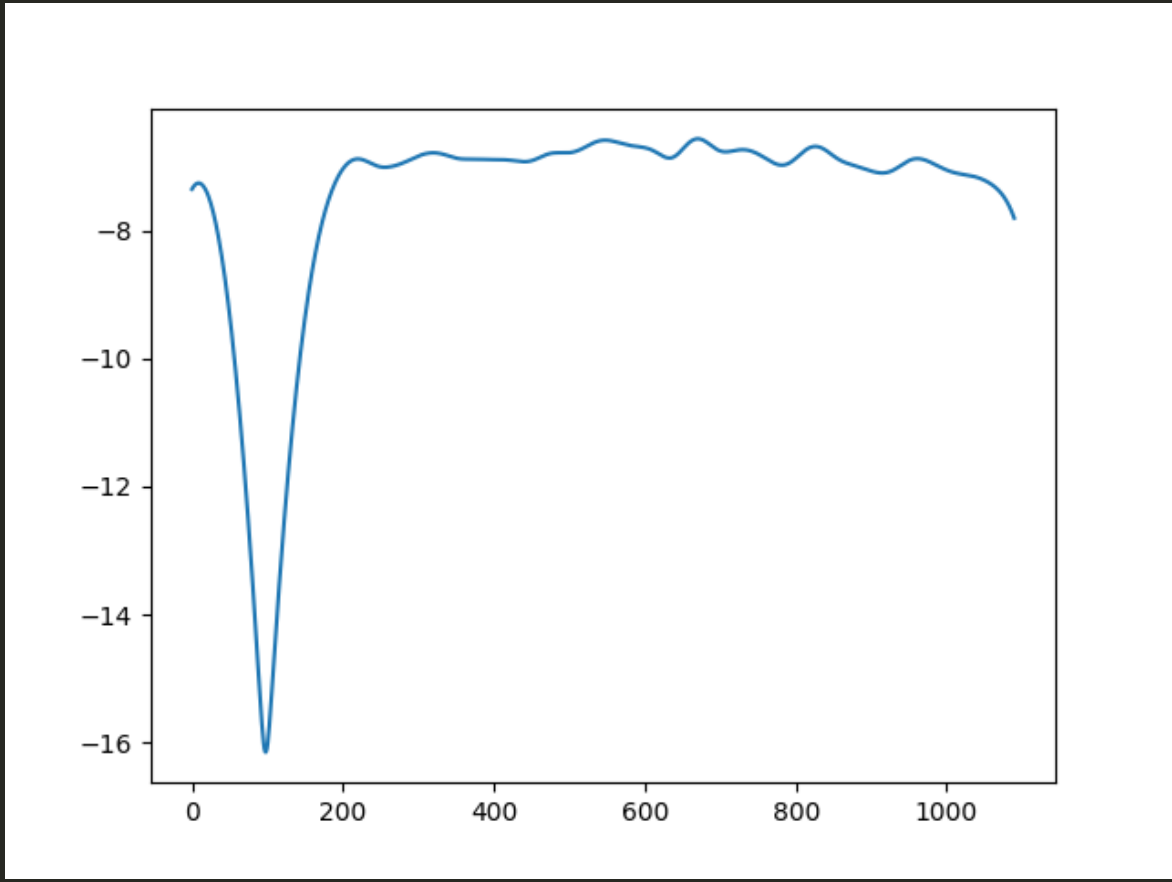
1. Assert that seeds must not be implausible - Removed insufficiently many seeds
    - Too close to the ends of the pod
    - Too large given pod dimensions
2. Real z position of seeds of a pod is a sample from some probability distribution, fit and paramterize the distribution to classify seeds.
    - 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)

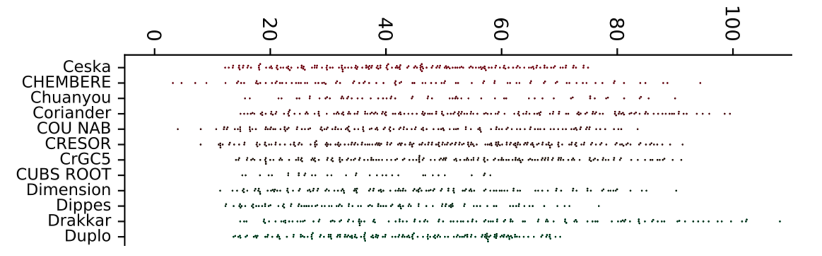
- Beak has no Real Seeds™ and low density
- Expect a gap in real z of detected seeds

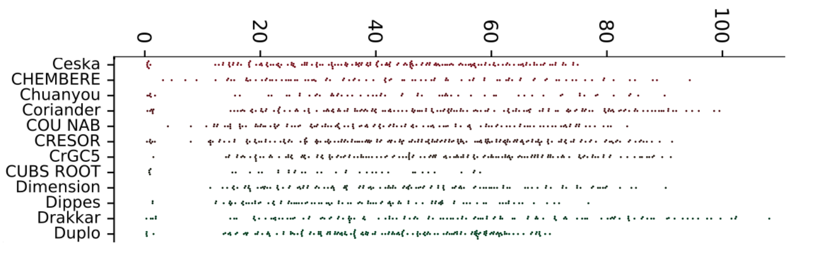


- Use KDE to find density of seeds as function of real z



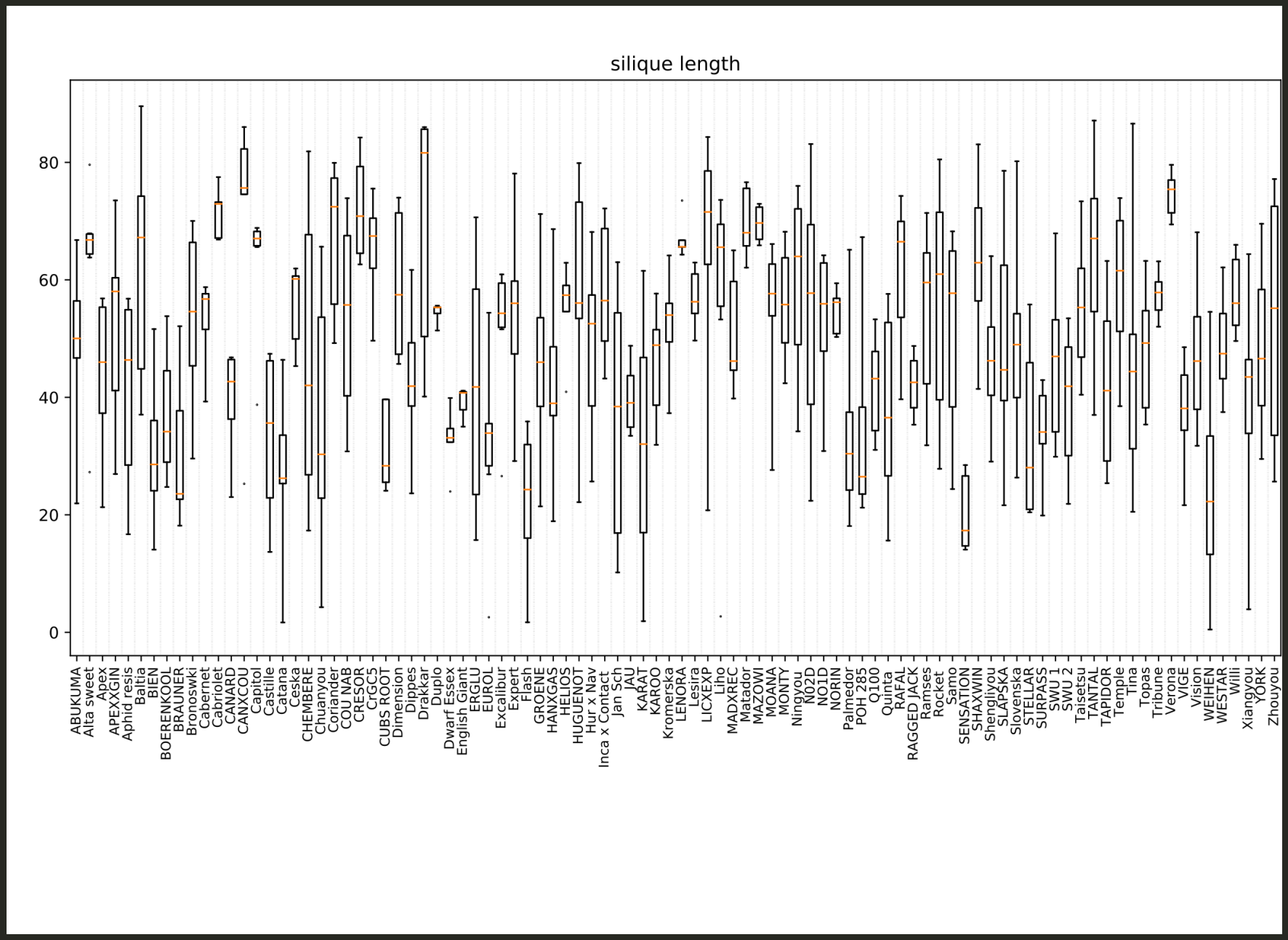
- First seed has real z less than 100?
- Find the local minimum at lowest real z where log(KDE)<-10
- Keep seeds with greater real z
- Profit

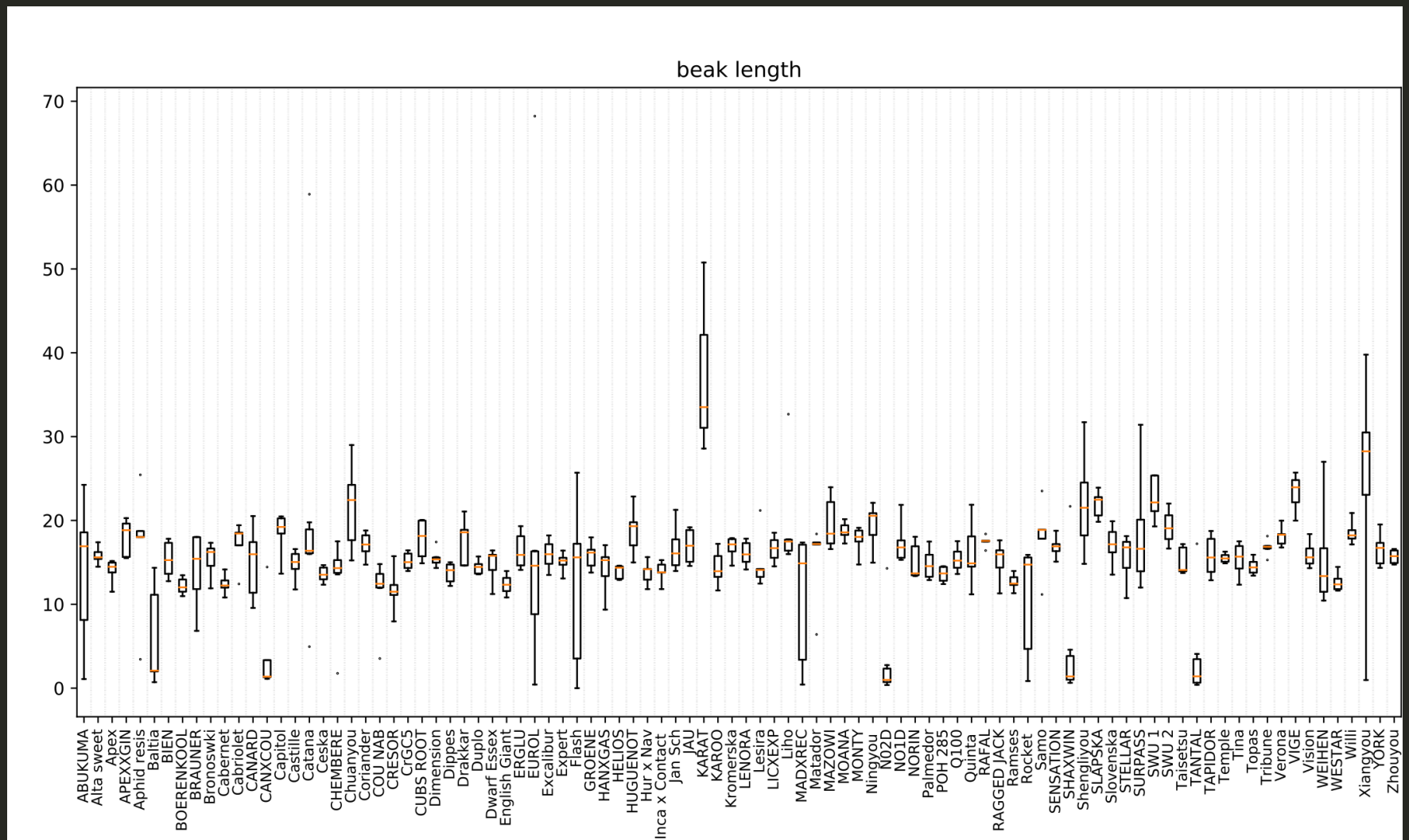

0
20
40
60
80
100
Ceska
CHEMBERE
Chuanyou
Coriander
COU NAB
CRESOR
CrGC5
CUBS ROOT
Dimension
Dippes
Drakkar
Duplo


0
20
40
60
80
100
Ceska
CHEMBERE
Chuanyou
Coriander
COU NAB
CRESOR
CrGC5
CUBS ROOT
Dimension
Dippes
Drakkar
Duplo

Beak and Silique length

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


silique length
0
20
40
60
80
ABUKUMA
Alta sweet
Apex
APEXXGIN
Aphid resis
Baltia
BIEN
BOERENKOOL
BRAUNER
Bronoswki
Cabernet
Cabriolet
CANARD
CANXCOU
Capitol
Castille
Catana
Ceska
CHEMBERE
Chuanyou
Coriander
COU NAB
CRESOR
CrGC5
CUBS ROOT
Dimension
Dippes
Drakkar
Duplo
Dwarf Essex
English Giant
ERGLU
EUROL
Excalibur
Expert
Flash
GROENE
HANXGAS
HELIOS
HUGUENOT
Hur x Nav
Inca x Contact
Jan Sch
JAU
KARAT
KAROO
Kromerska
LENORA
Lesira
LICXEXP
Liho
MADXREC
Matador
MAZOWI
MOANA
MONTY
Ningyou
N02D
NO1D
NORIN
Palmedor
POH 285
Q100
Quinta
RAFAL
RAGGED JACK
Ramses
Rocket
Samo
SENSATION
SHAXWIN
Shengliyou
SLAPSKA
Slovenska
STELLAR
SURPASS
SWU 1
SWU 2
Taisetsu
TANTAL
TAPIDOR
Temple
Tina
Topas
Tribune
Verona
VIGE
Vision
WEIHEN
WESTAR
Willi
Xiangyou
YORK
Zhouyou


beak length
70
60
50
40
30
20
10
0
ABUKUMA
Alta sweet
Apex
APEXXGIN
Aphid resis
Baltia
BIEN
BOERENKOOL
BRAUNER
Bronoswki
Cabernet
Cabriolet
CANARD
CANXCOU
Capitol
Castille
Catana
Ceska
CHEMBERE
Chuanyou
Coriander
COU NAB
CRESOR
CrGC5
CUBS ROOT
Dimension
Dippes
Drakkar
Duplo
Dwarf Essex
English Giant
ERGLU
EUROL
Excalibur
Expert
Flash
GROENE
HANXGAS
HELIOS
HUGUENOT
Hur x Nav
Inca x Contact
Jan Sch
JAU
KARAT
KAROO
Kromerska
LENORA
Lesira
LICXEXP
Liho
MADXREC
Matador
MAZOWI
MOANA
MONTY
Ningyou
N02D
NO1D
NORIN
Palmedor
POH 285
Q100
Quinta
RAFAL
RAGGED JACK
Ramses
Rocket
Samo
SENSATION
SHAXWIN
Shengliyou
SLAPSKA
Slovenska
STELLAR
SURPASS
SWU 1
SWU 2
Taisetsu
TANTAL
TAPIDOR
Temple
Tina
Topas
Tribune
Verona
VIGE
Vision
WEIHEN
WESTAR
Willi
Xiangyou
YORK
Zhouyou

Resources

Git Repository [1]

This Presentation [2]

Numpy [3]

Matplotlib [4]

Seaborne [5]

---

1 https://github.com/NPPC-UK/ct_scanner_plotting

2 https://git.friedersdorff.com/max/ct_plotting.git

3 https://www.numpy.org/
4 https://matplotlib.org/
5 https://seaborn.pydata.org/