

Summary exploratory analysis Pear Data

Sabine K. Schnabel, Biometris

July 2, 2015

General information about the data sets

Meta data

- It is not yet for sure in how far we will use the meta data file at all.
- Excel table with information about the different orchards
- There are 18 orchards with partial information in the data set.
- The variables (with different degrees of information and variation):
boomvorm, boomleeftijd, onderstam, grondsoort, fractie afslibbaar, fertigatie, voeding N, voeding P, voeding K, wortelsnoei, behang (%), commerciële pluk, pluk proef, CO₂ prod.(ml/kg.uur), ethyleenproductie (l/kg*h), %droge stof, vrucht-Ca, vrucht-K, vrucht-K/Ca, vrucht-P, vrucht-Mg, vrucht-N.

Large data set

- Excel table with information on all measurement and indicators taken on the pears
- The table has the following information on fruit level:
 - Basic information about the record:
regel, vrucht nr, herkomst, product, batch, duplo, meetdatum, SmartFresh, destructief/non-destructief, meting
 - General measurement (repeated at different time points):
kleurindex (1=groen-5=geel), **penetrometerwaarde/firmness**

- Measurements from the Pigment Analyzer:
NDVI -bloskant, NAI-bloskant, NDVI - groene zijde, NAI-groene zijde, NDVI-gemiddeld, NAI-gemiddeld
 - Measurements from the DA-meter (non destructive chlorophyll-analyzer):
DA-index bloskant/red, DA-index groene zijde/green, DA-index gemiddeld, DA-indexblos-groen
 - Measurements from the Aweta AFS Acoustic firmness sensor:
Stevigheid AFS-bloskant, M (g) bloskant, F0 bloskant, Stevigheid AFS-groene zijde, M (g) groene zijde, F0 groene zijde, Stevigheid AFS-gemiddeld, M (g) AFS % verschil, F0 AFS-gemiddeld
 - Underwater weighing to determine density: vruchtgewicht(g), dichtheid(g/ml)
 - % droge stof, TSS (Brix), zetmeelindex (1-10)
 - Schilvlekjes-score, Schilvlekjes index
 - China simulatie: rot, zwarte stelen index, hol&bruin-score
- Additional there is substantial information about LBA and Hue strata as well as a large amount of metabolomics measurements. The latter are at batch level while the former are on fruit level.
 - A lot of the measurements are non-destructive (that includes the LBA/Hue measurements), but the most meaningful measure **firmness** is destructive and also metabolomics and other measurements (**dry matter** etc.) result in the loss of the fruit.

Descriptive analysis of the data

Meta data

Since this data set is rather small and most information are only partial or do not contain enough variation, no further descriptive analysis was done.

Large data set

These data include information on Conference pears as well as Elstar apples. For the sake of this analysis I excluded the apples from the data set. The resulting data set has 8720 records and includes numerous columns (with basic information and phenotypic traits). To my understanding there are

4320 different pears in the data set. Since the most important measurement is destructive at every measurement for firmness the measured fruits are sacrificed. The fruits are uniquely identified through the fruit number, the orchard indication, the batch and duplo.