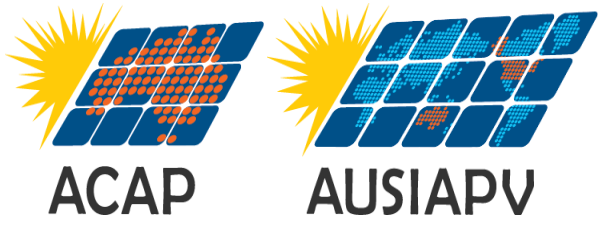
Interpreting Manufacturing Variance using a Data Mining Approach



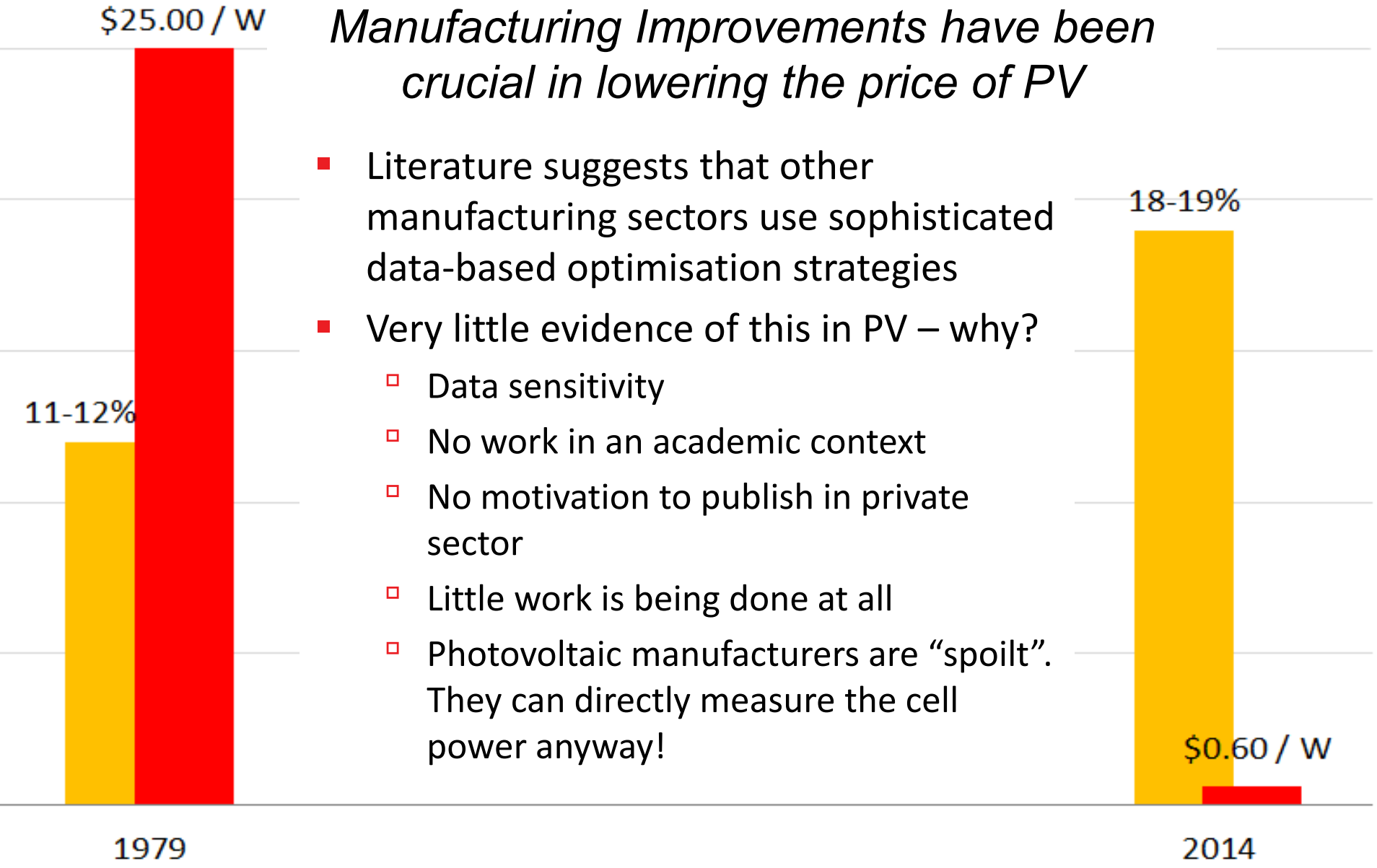
Oct 2014

Rhett Evans, Jonathon Dore,

Erik Van Voorthuysen, Jingbing Zhu, Martin Green.

evansrhett@gmail.com

Why Manufacturing Research?



Manufacturing Improvements have been crucial in lowering the price of PV

- Literature suggests that other manufacturing sectors use sophisticated data-based optimisation strategies
- Very little evidence of this in PV – why?
  - Data sensitivity
  - No work in an academic context
  - No motivation to publish in private sector
  - Little work is being done at all
  - Photovoltaic manufacturers are “spoilt”. They can directly measure the cell power anyway!

Details of this Study

Materials

1. Data from mc-Si cells from Suntech in Wuxi, China, 2013
2. Typical performance parameters for 20,000 cells
3. Data from two lines with different screen print patterns.

Methods

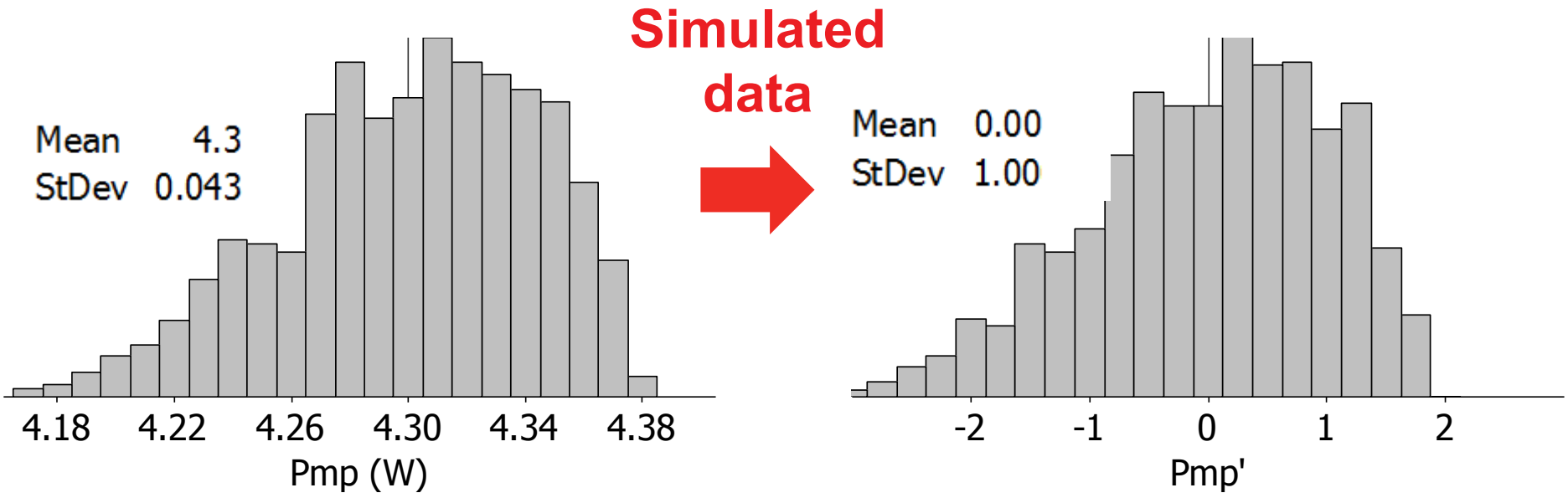
1. Data Normalisation
2. Simple data mining / data manipulation
3. Data Rotation to isolate variance

Aim

1. Demonstration of statistical techniques
2. Improved understanding of production variance
3. Improved understanding of cell operation

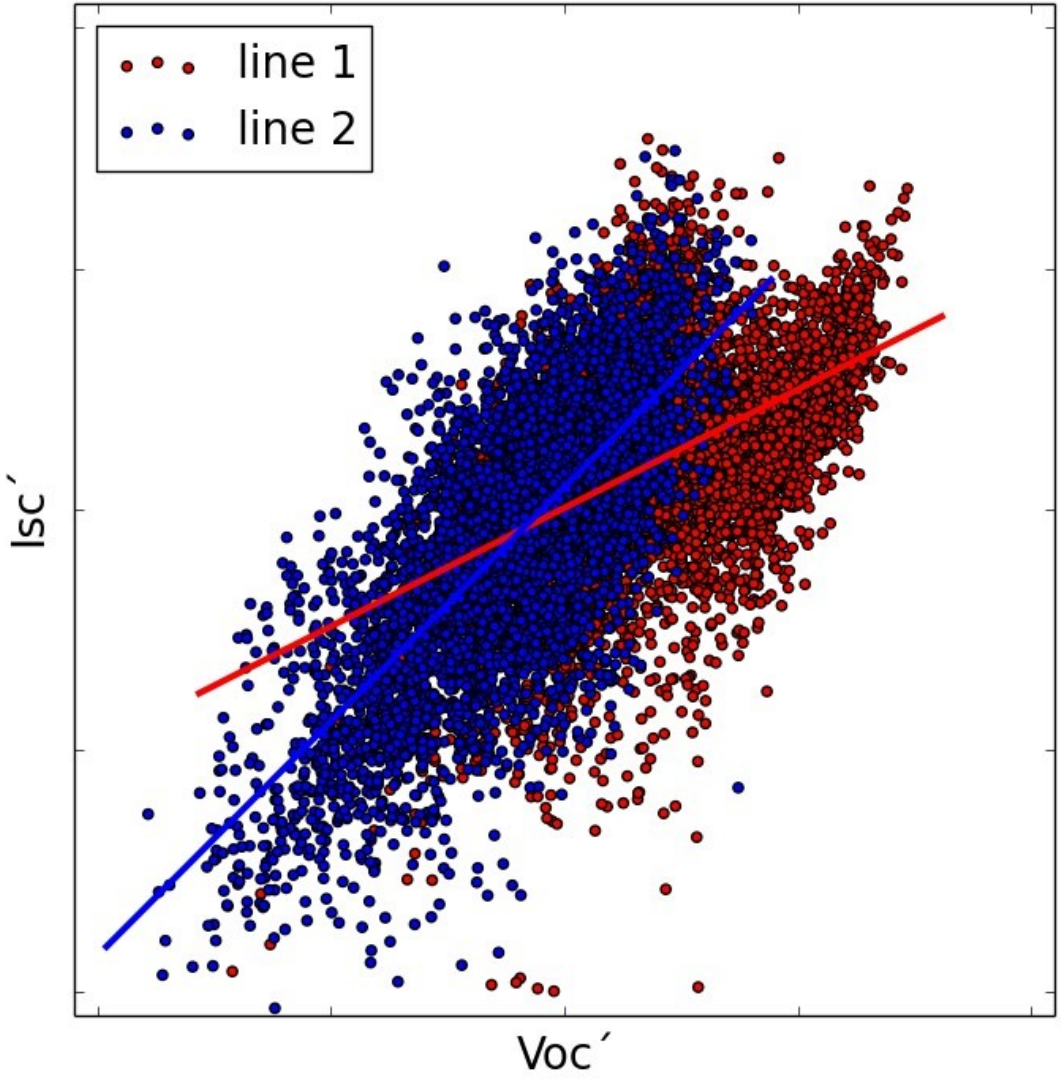
Data Normalisation

- All data must be normalised to share it publically
- Data normalisation most often means ***Mean 0, Variance 1***



- Here, variance in FF, Voc and Isc is scaled ***relative*** to Pmp
  - E.g 0.043/4.3 = 1% would be considered 1 unit of variance in Voc, Isc and FF.

Isc vs Voc data

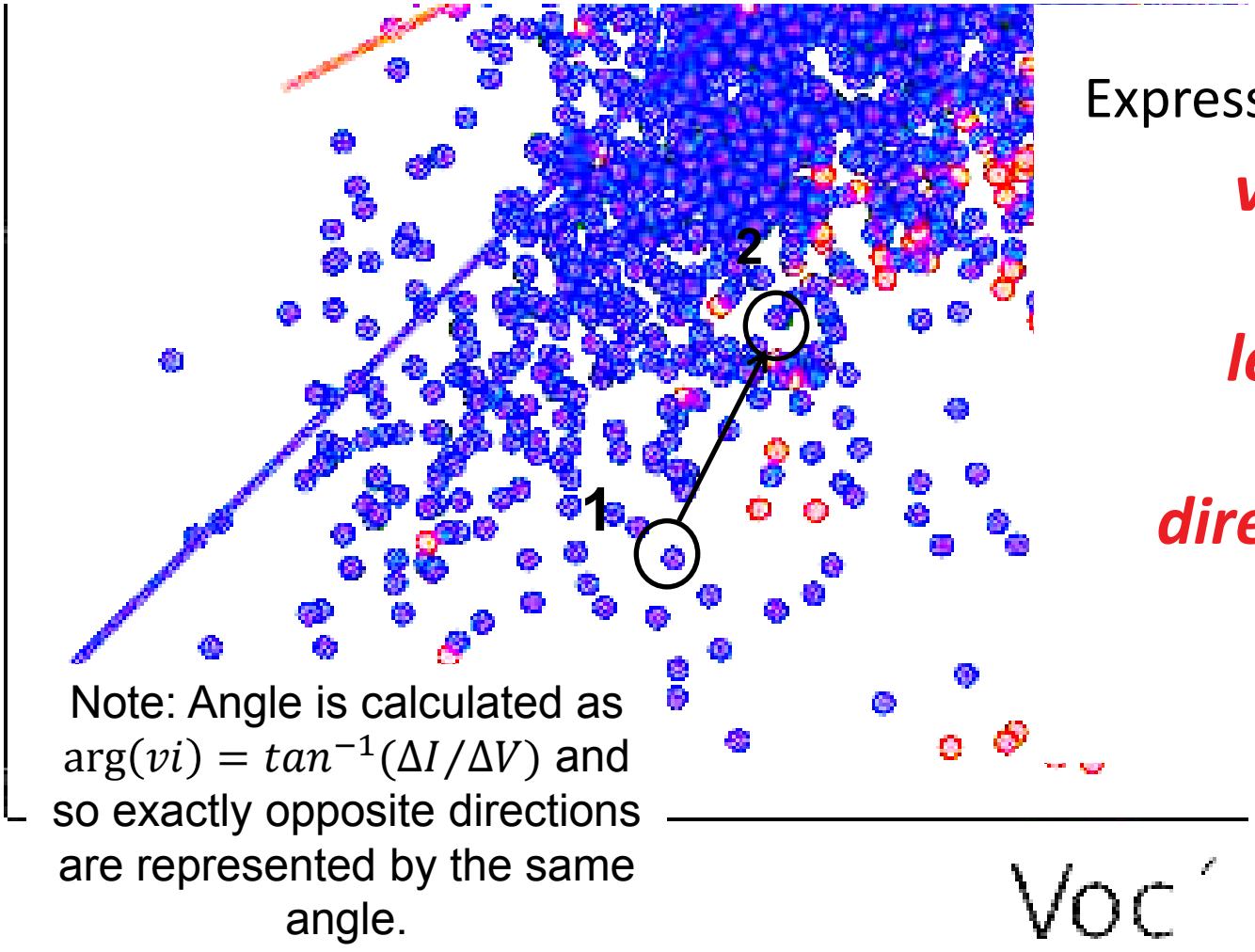


Data is from
20,000 cells

There is a different Isc' vs Voc' trend in both data sets with a different line of best fit.

Isc vs Voc data

1. Consider some point on this relationship in time
2. Look at what changes to make the next cell



Express this change as a

vector $\overrightarrow{vi}$,

with

length $\left|\overrightarrow{vi}\right|$

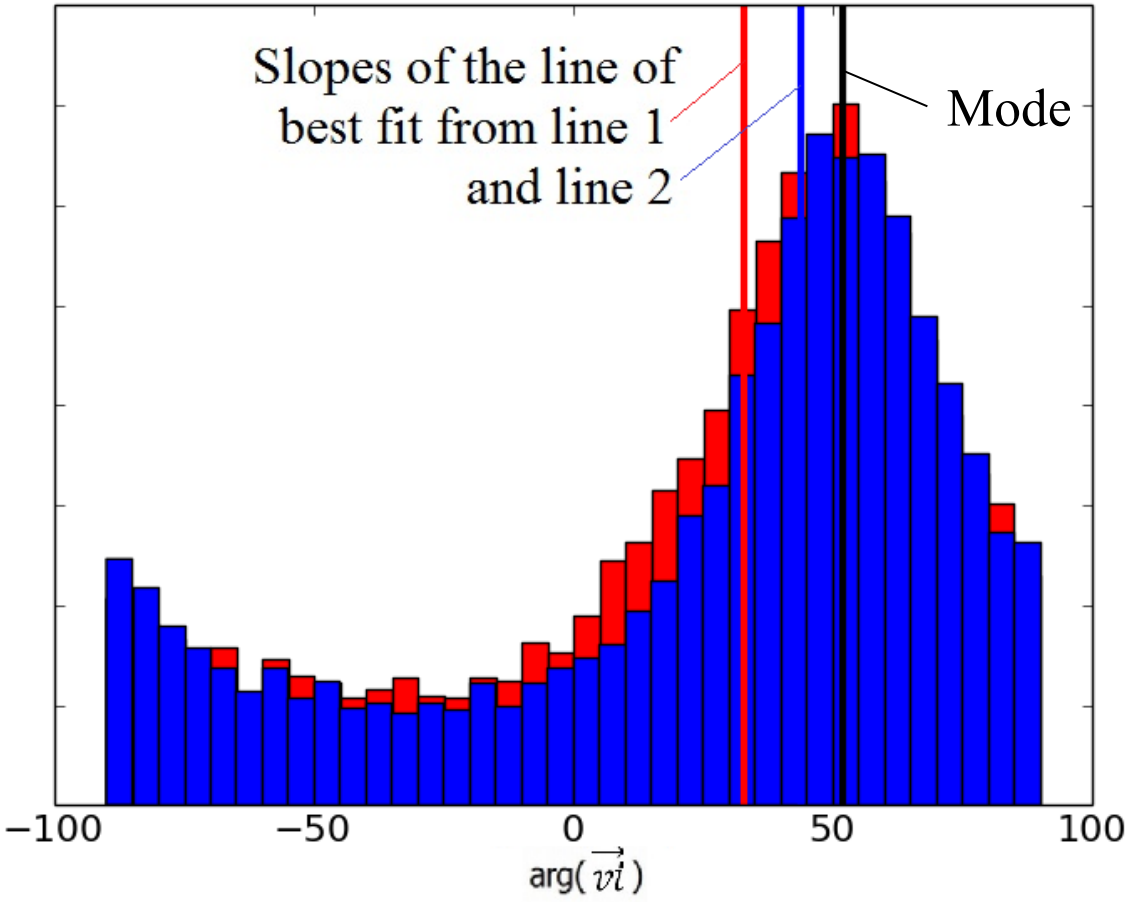
and

direction $\arg\left(\overrightarrow{vi}\right)$

Note: Angle is calculated as $\arg(vi) = tan^{-1}(\Delta I/\Delta V)$ and so exactly opposite directions are represented by the same angle.

Isc vs Voc data

- Plot the histogram of the direction angle of this vector
- The modal response is 52^{o} in both cases
- The lines of best fit have lower slope

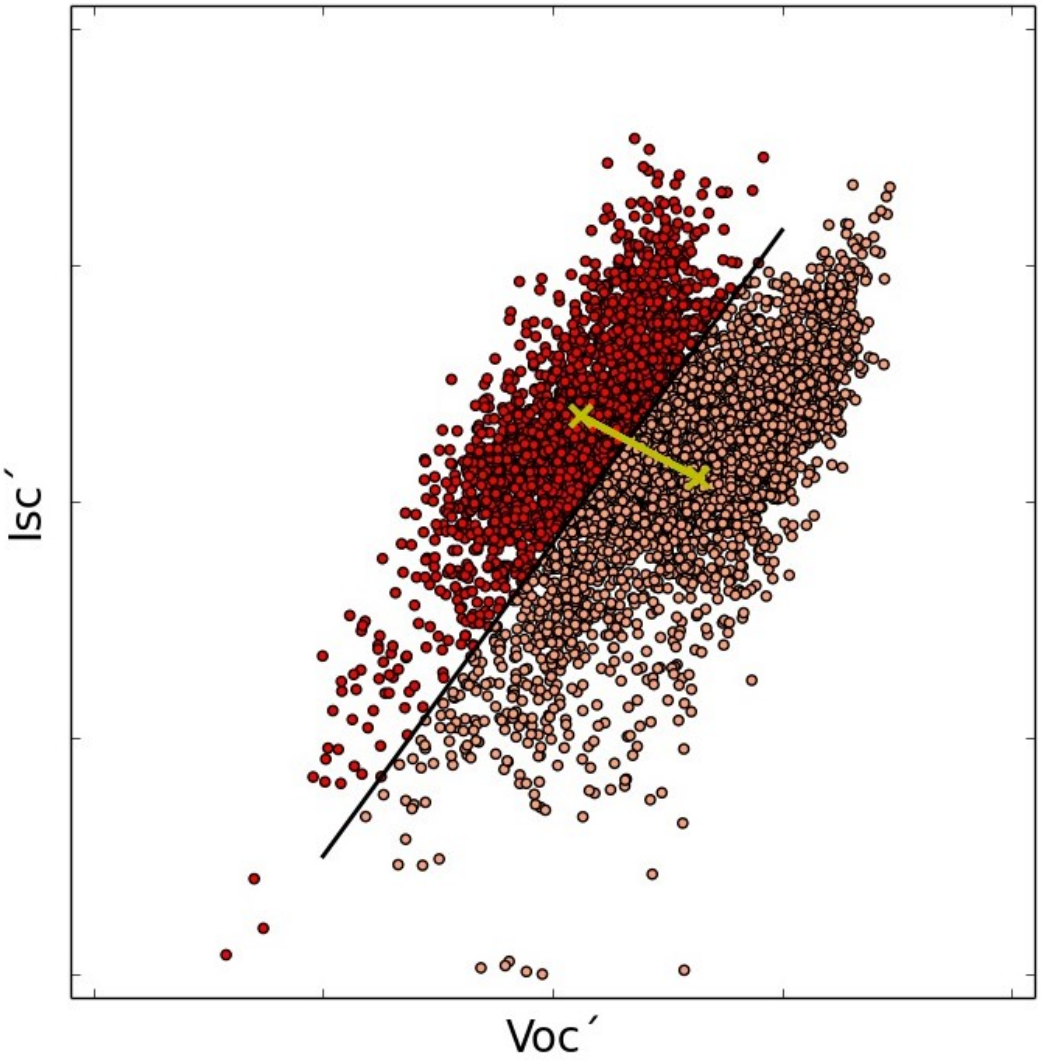


PC1D, can model this sort of Voc/Isc relationship by ***varying the lifetime*** of the wafer.

The exact slope is providing information on the recombination properties

Isc vs Voc data

- What about the double distribution in the line 1 data?

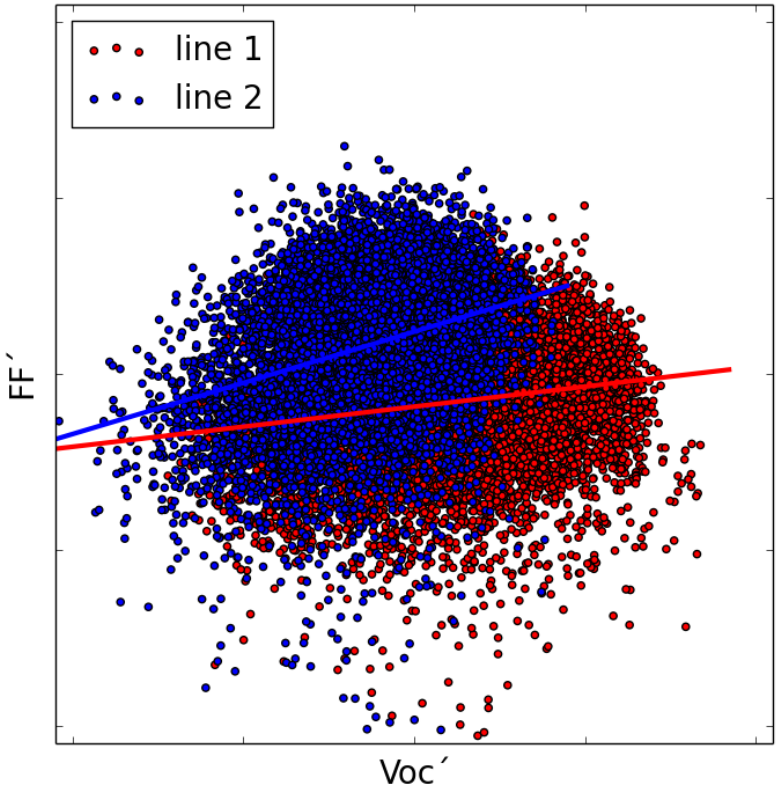


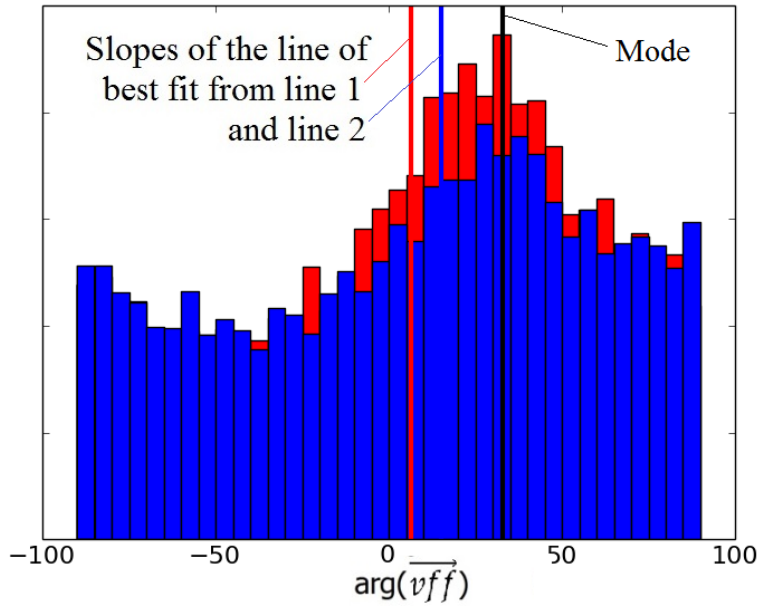
- The data can be split approximately as shown
- The centroid of each data-set calculated and a line plotted between them.
- The angle of this line is between -25 and -30

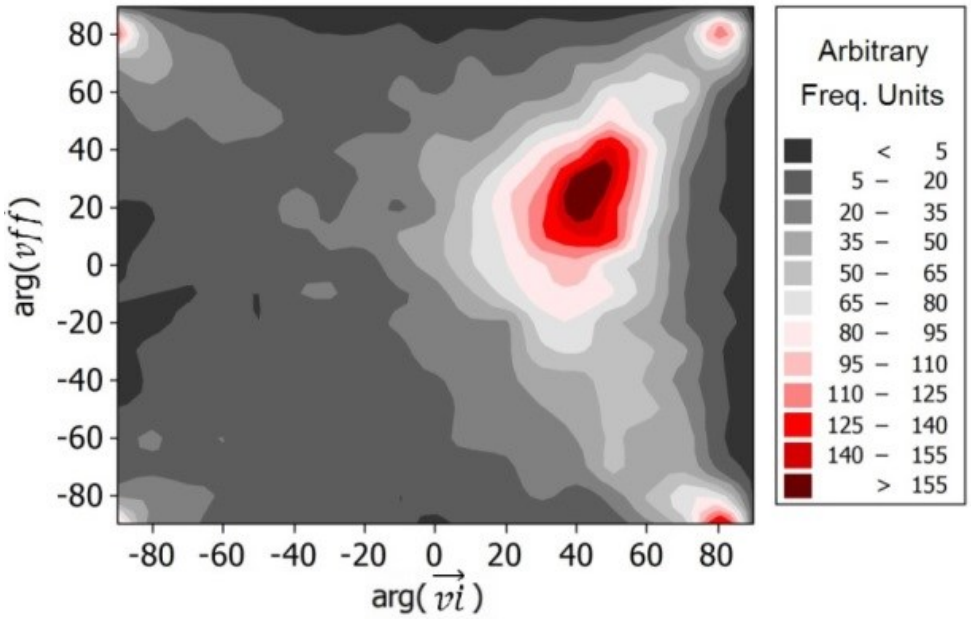
PC1D models say this is the Voc / Isc relationship expected from ***wafer resistivity changes***

FF vs Voc data

- The same vector analysis can be done for FF / Voc



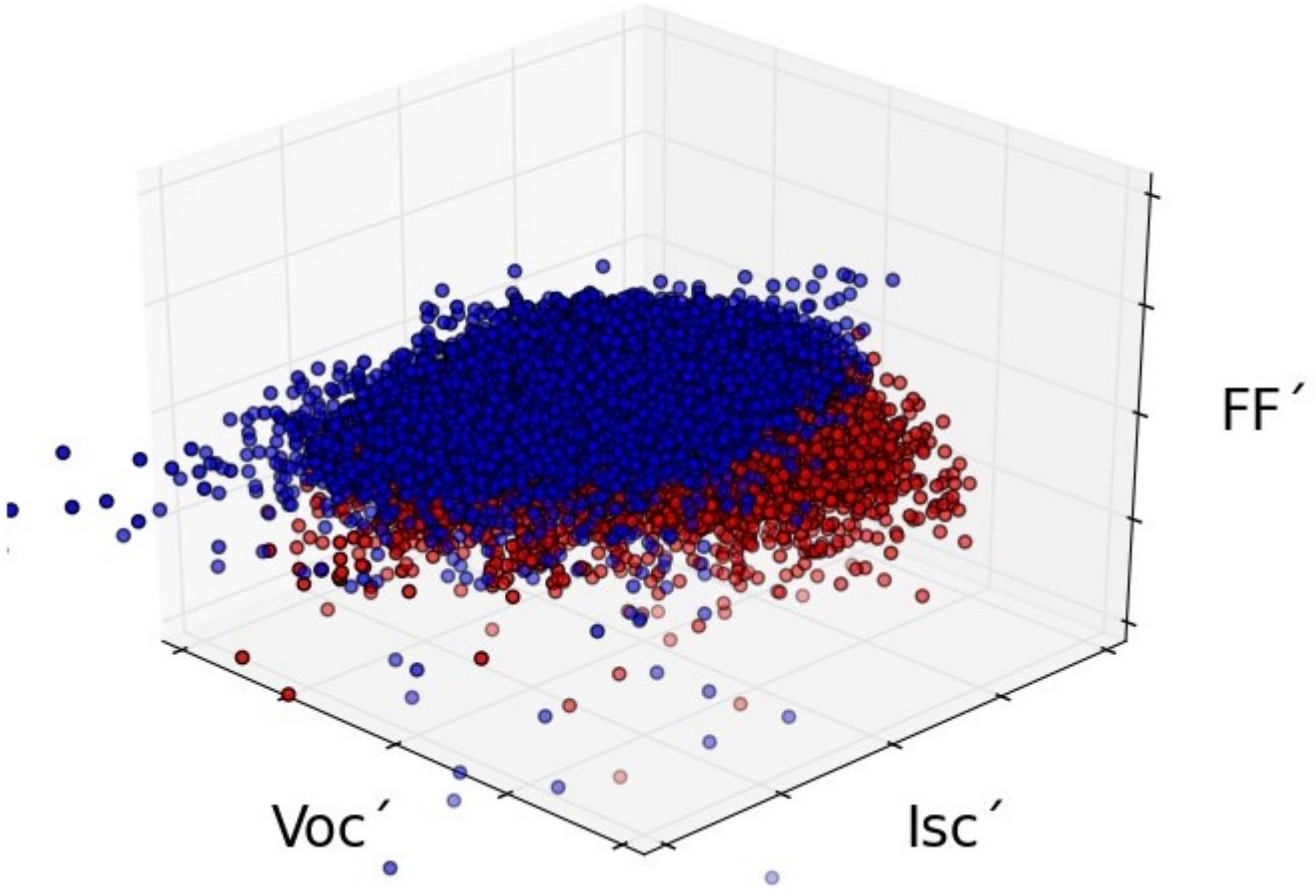




- Plotting together suggests a common causality.

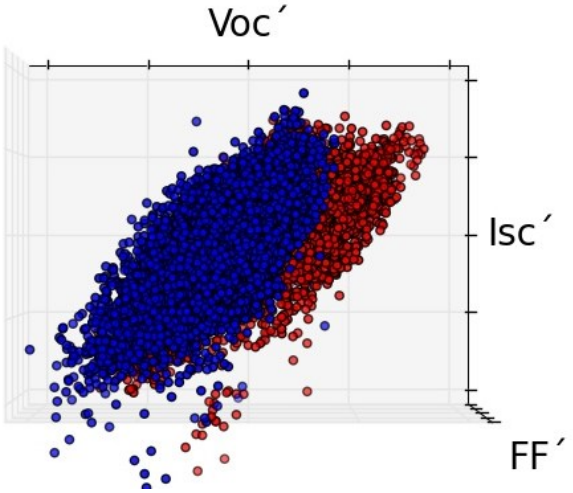
Voc / Isc / FF data in 3D space

- Consider the cell data in three-dimensional space with wafer lifetime a dominant source of variation

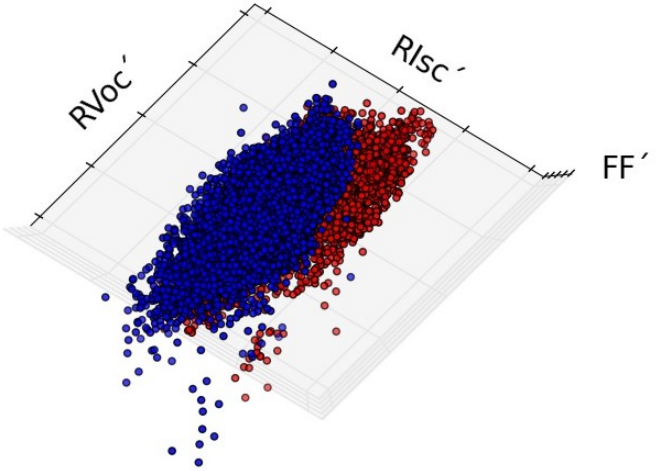


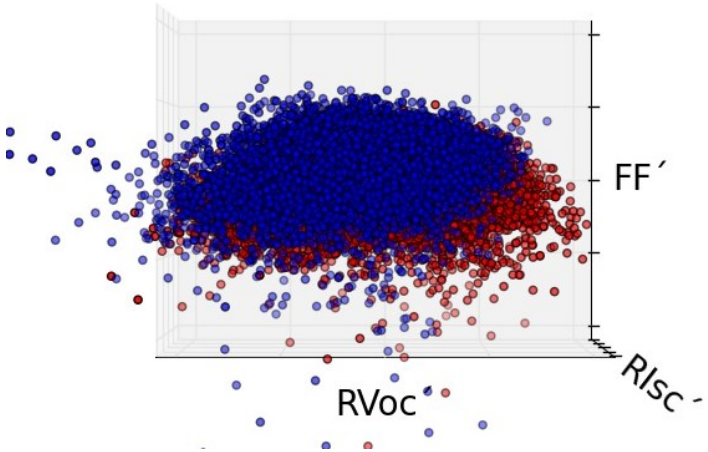
Axis Rotation

- It is possible to rotate the data axis wrt the cell data to isolate the effects of material variation.

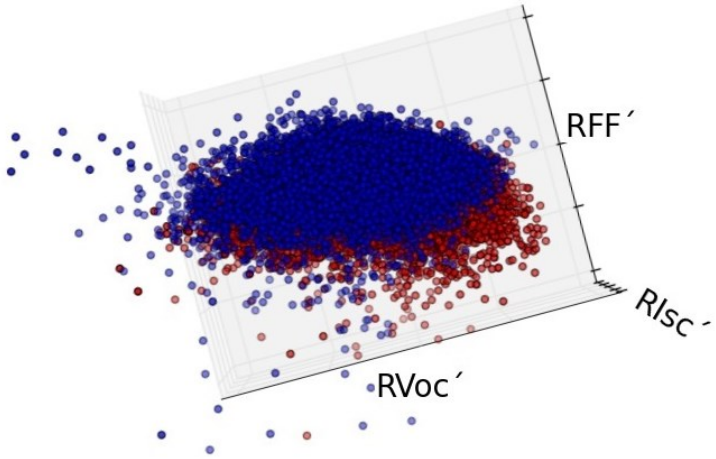


Rotate Voc / Isc Axis

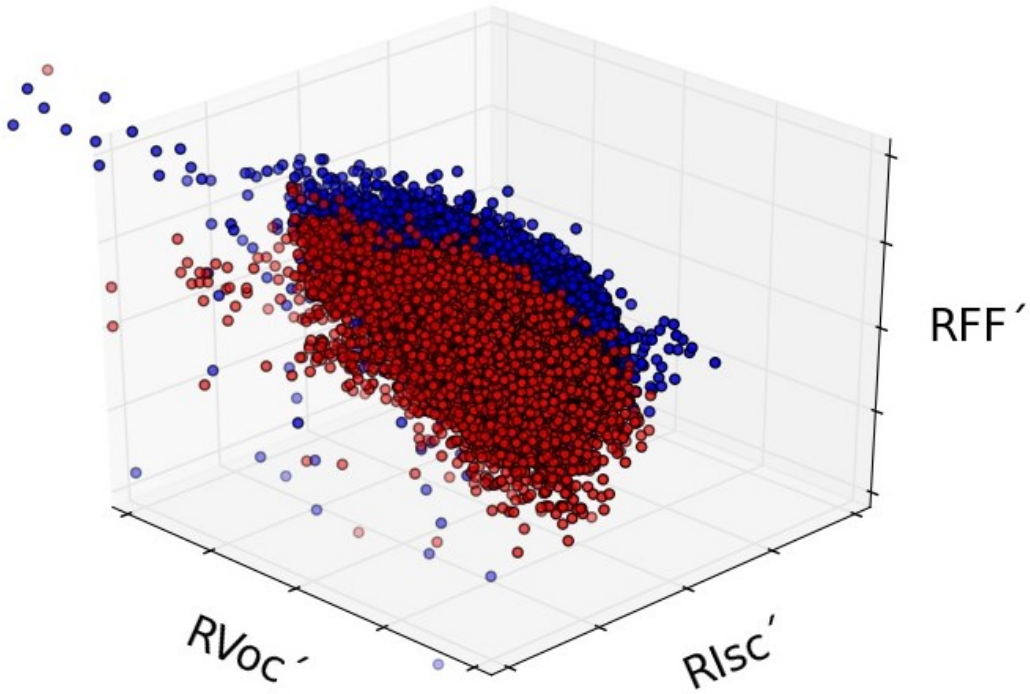




Rotate FF / Voc Axis



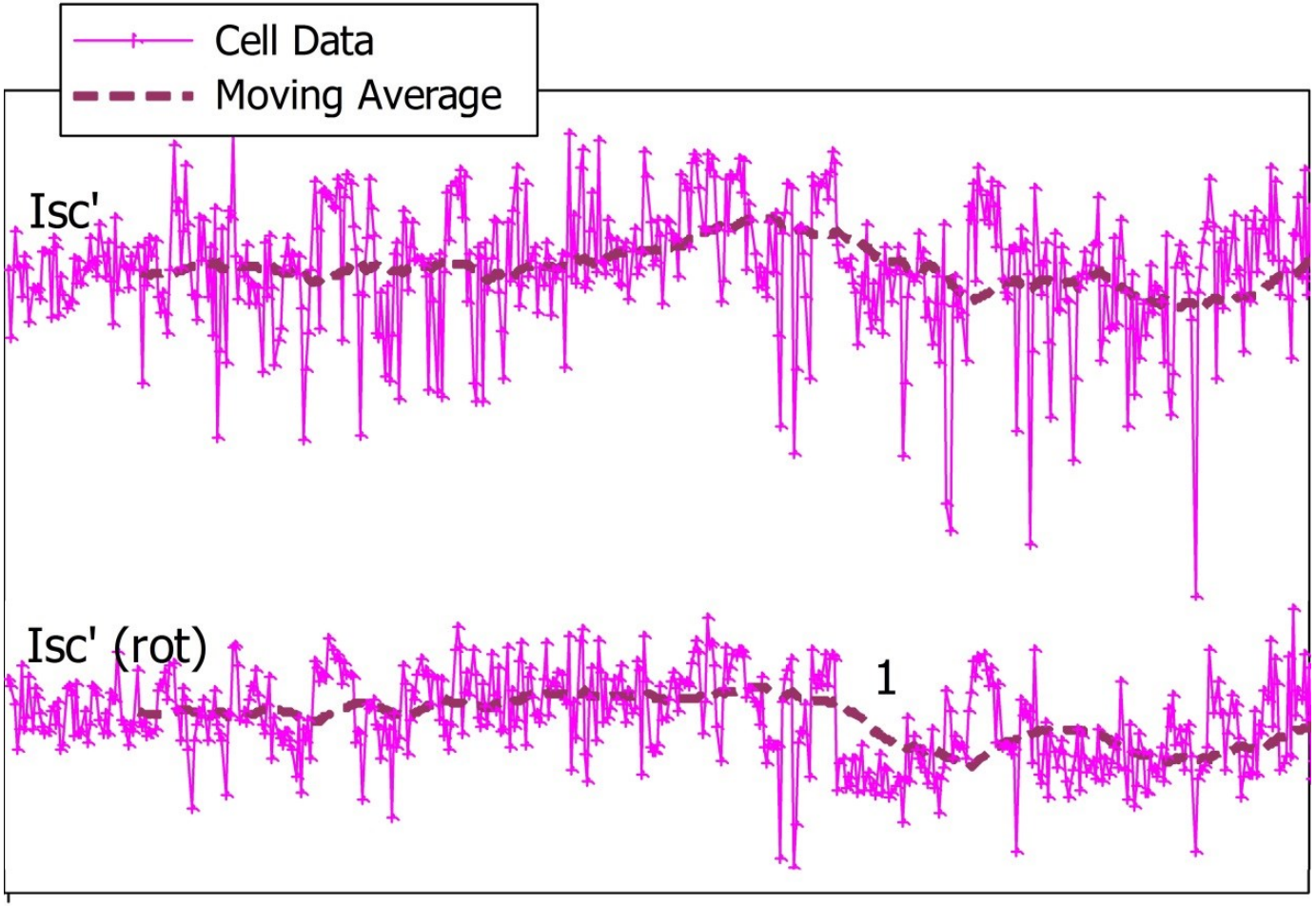
Interpreting Rotated Data



- RVoc’ axis now represents material lifetime variation
- RIsc’ axis describes variation in Isc’ unrelated to material
- RFF’ axis describes variation in FF’ unrelated to material

Using Rotated Data

- Undertaking SPC on rotated components means material effects are removed and other process faults are highlighted



Summary and Further Work

Summary

- Data optimisation is an important part of future PV manufacturing
- Shown here are three basic and important multivariate approaches
  1. Data normalisation to share statistical techniques
  2. Some introductory data manipulation or data mining techniques
  3. Axis rotation to isolate variance
- Large data sets contain information on detailed cell properties

Future Work

- The precise angles in the Isc / Voc and FF / Voc relationships tell us about the detailed recombination properties of the cells.
  - This requires further investigation
- Build more extensive statistical models to characterise manufacturing.
  - Statistical models of optimal manufacturing size
  - Models describing optimal data sets for manufacturing