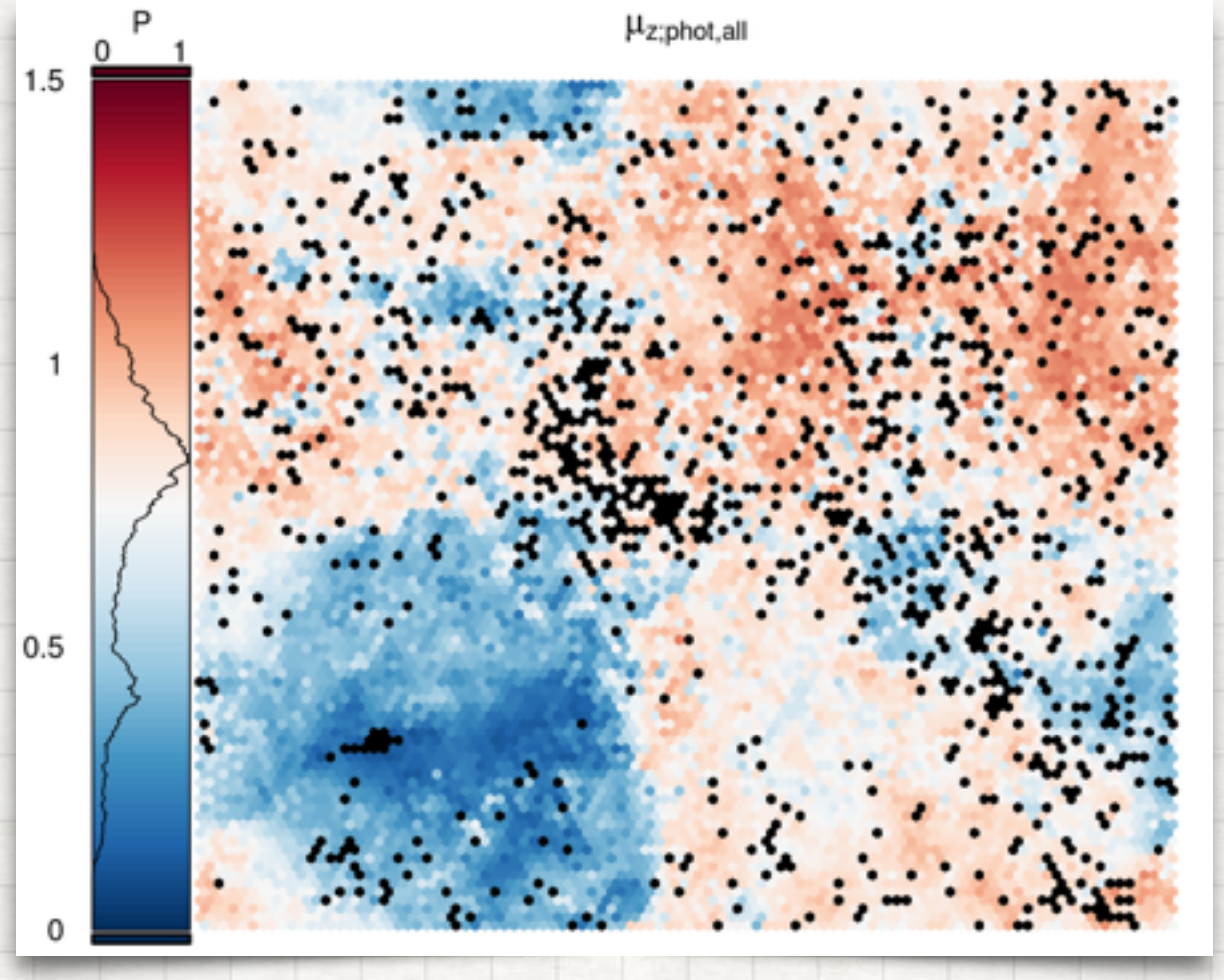
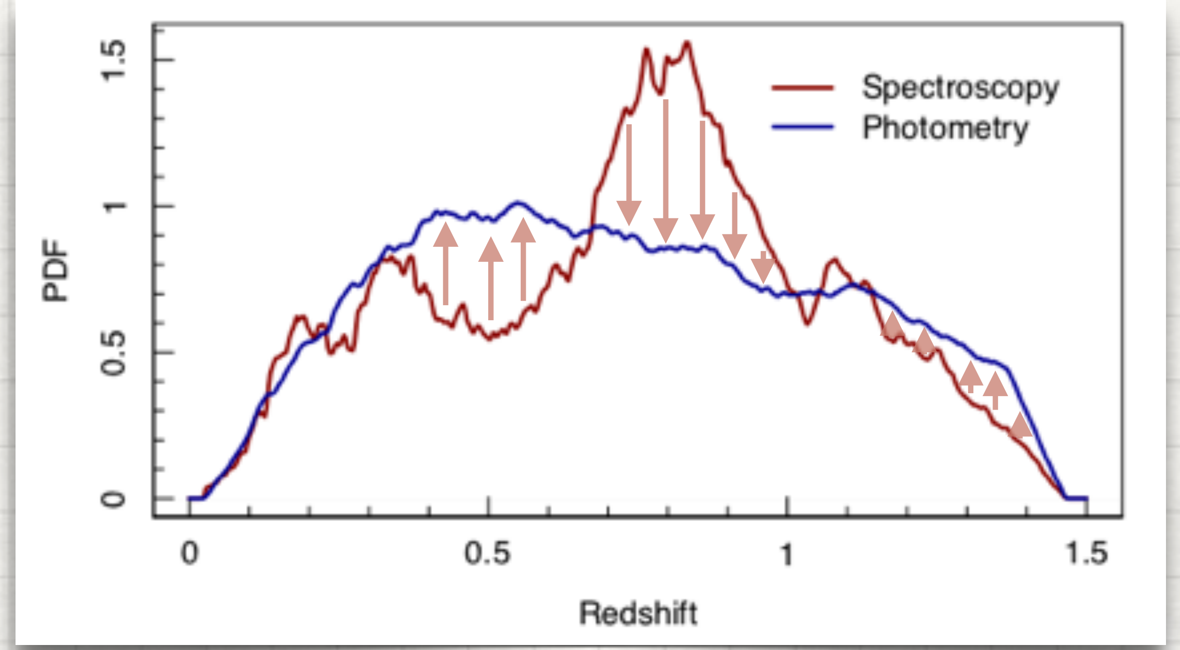
GCCL SEMINAR 06-03-2020

PHOTOMETRIC REDSHIFT CALIBRATION WITH SELF ORGANISING MAPS





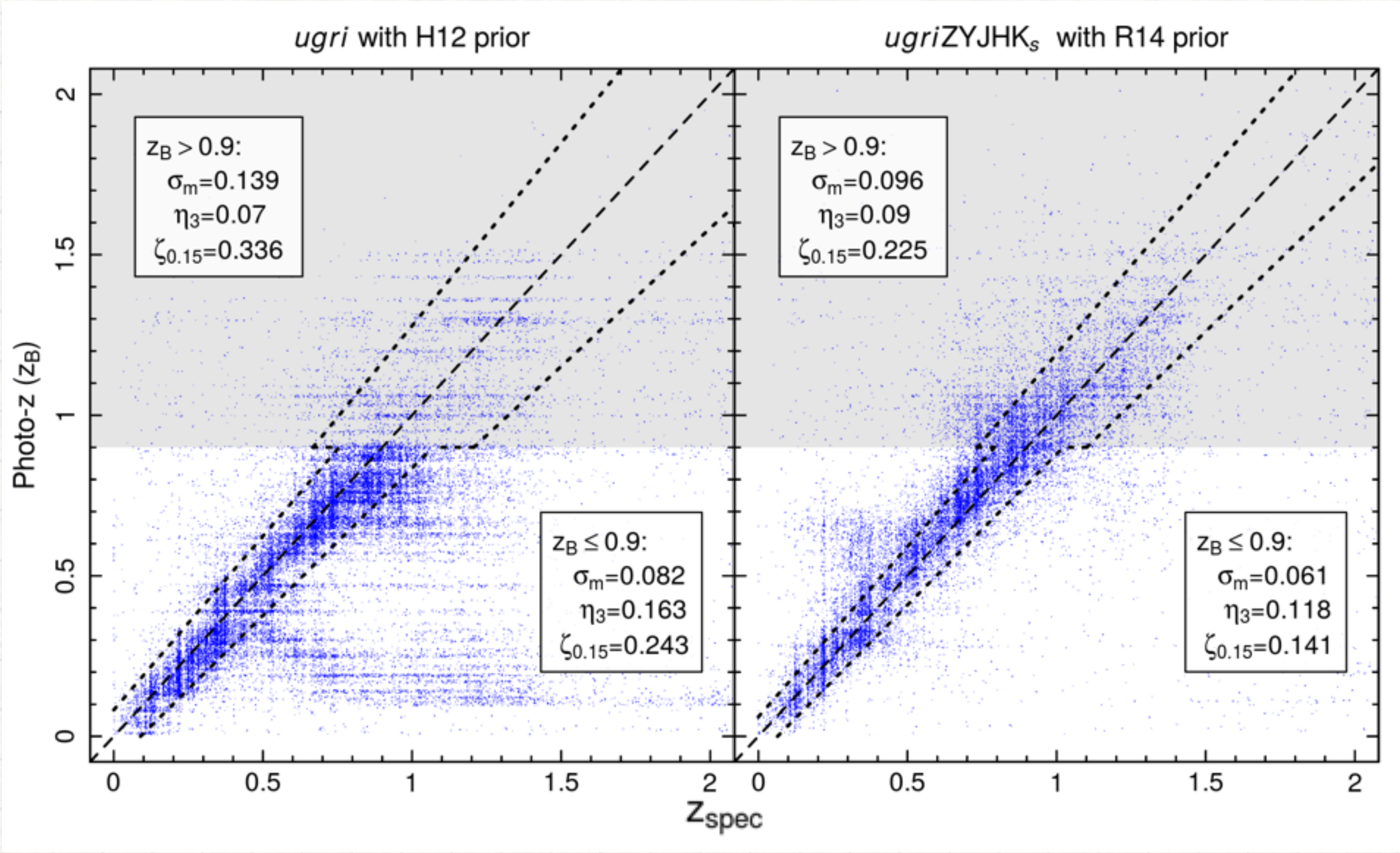
Angus H Wright
Research Fellow

German centre for cosmological lensing (GCCL)

Ruhr universität Bochum, Germany

REDSHIFT CALIBRATION FOR COSMIC SHEAR

PHOTOMETRIC REDSHIFTS



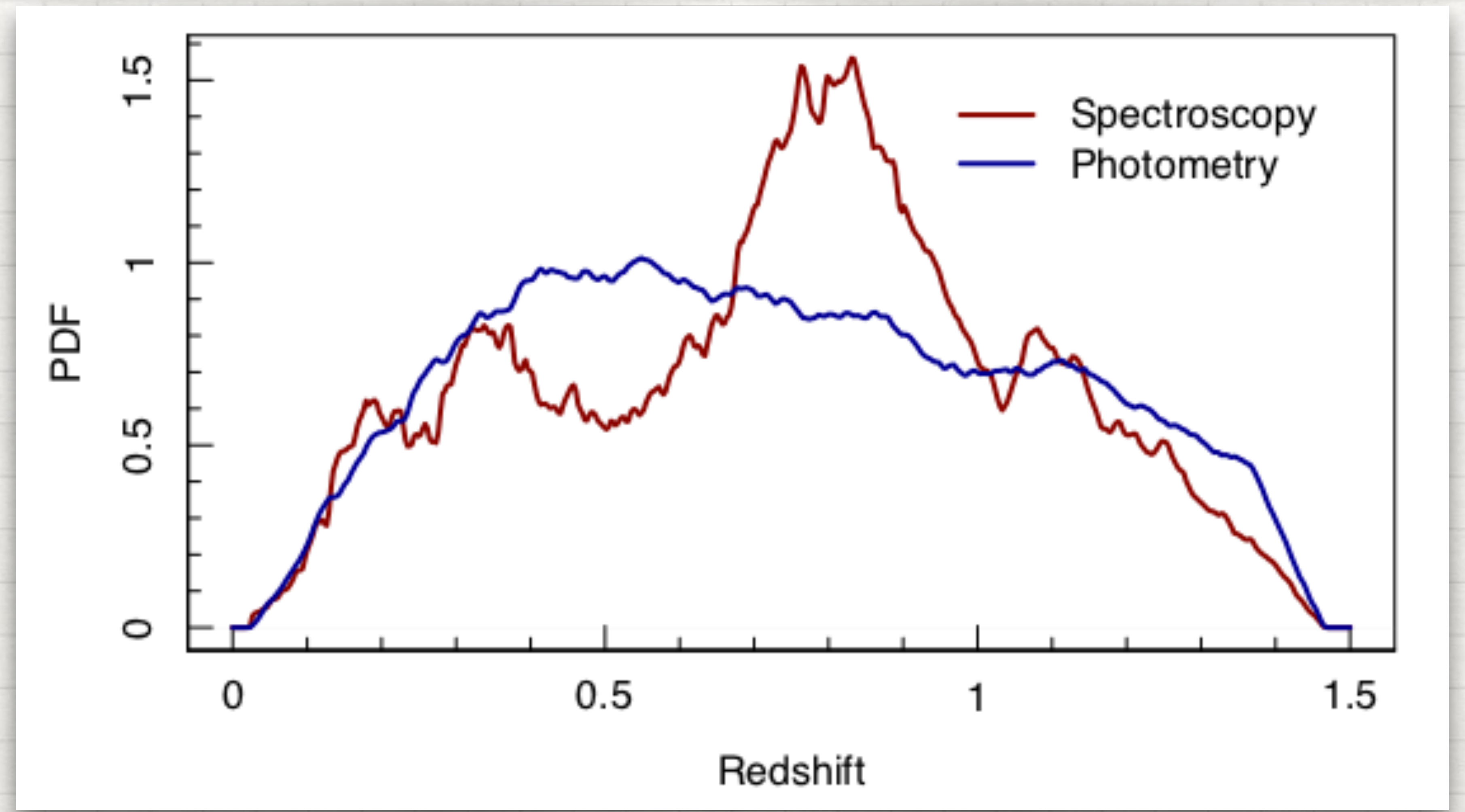
Photometric redshift point estimates correlate with, but are not the same as, true redshift.

↳ These are not sufficiently accurate for cosmic shear.

Wright et al (2019)

DIRECT REDSHIFT CALIBRATION FOR COSMIC SHEAR

SPECTROSCOPIC CALIBRATION



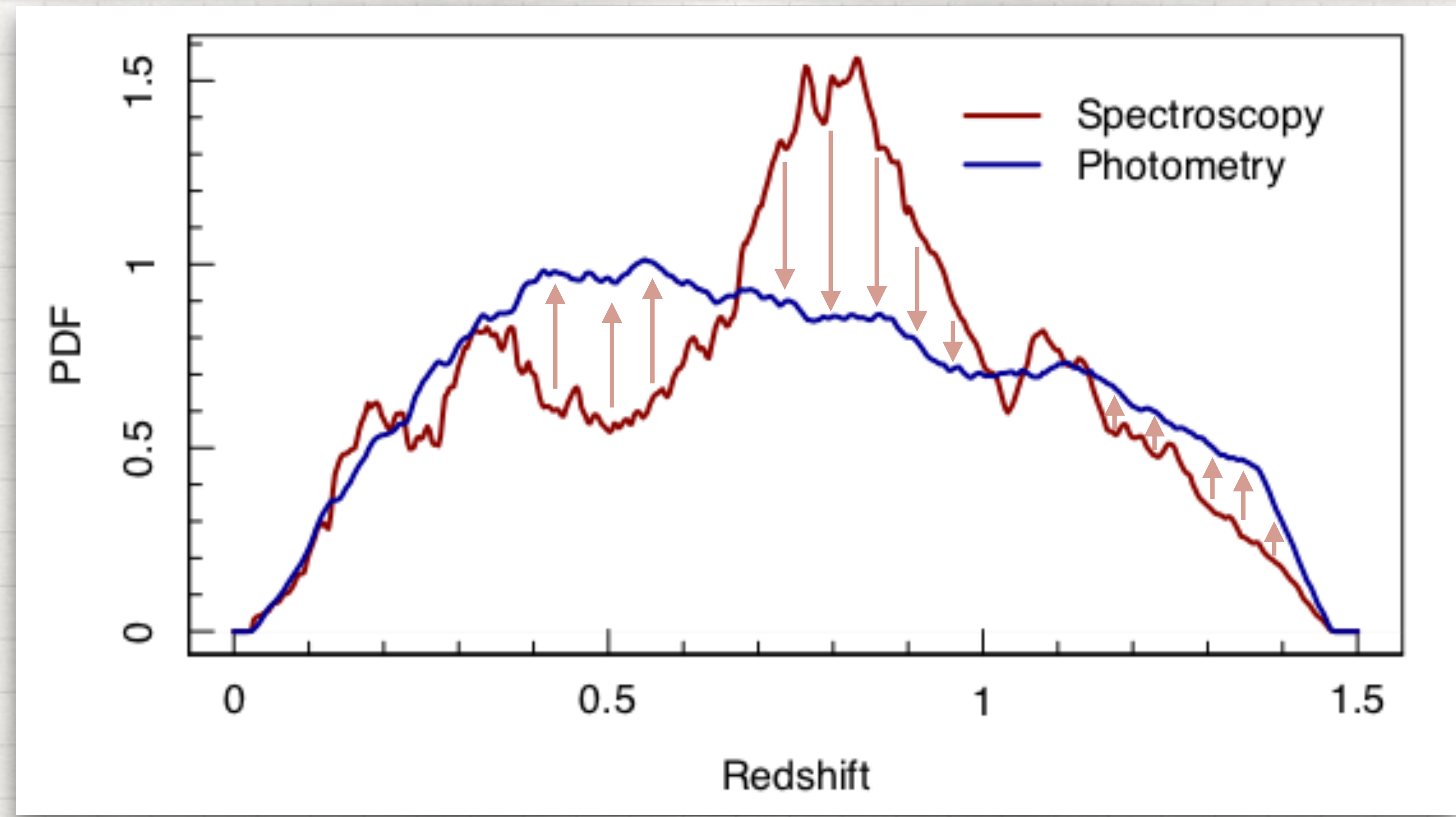
One calibration option is to use (high accuracy) spectroscopic redshifts to construct redshift distributions for the cosmic-shear (i.e. wide-field photometric) galaxies.

↳ Requires a method of mapping spectra onto wide-field galaxies

↳ Requires a representative spec-z sample *or* a method for determining which photometric sources are matched

DIRECT REDSHIFT CALIBRATION FOR COSMIC SHEAR

SPECTROSCOPIC CALIBRATION



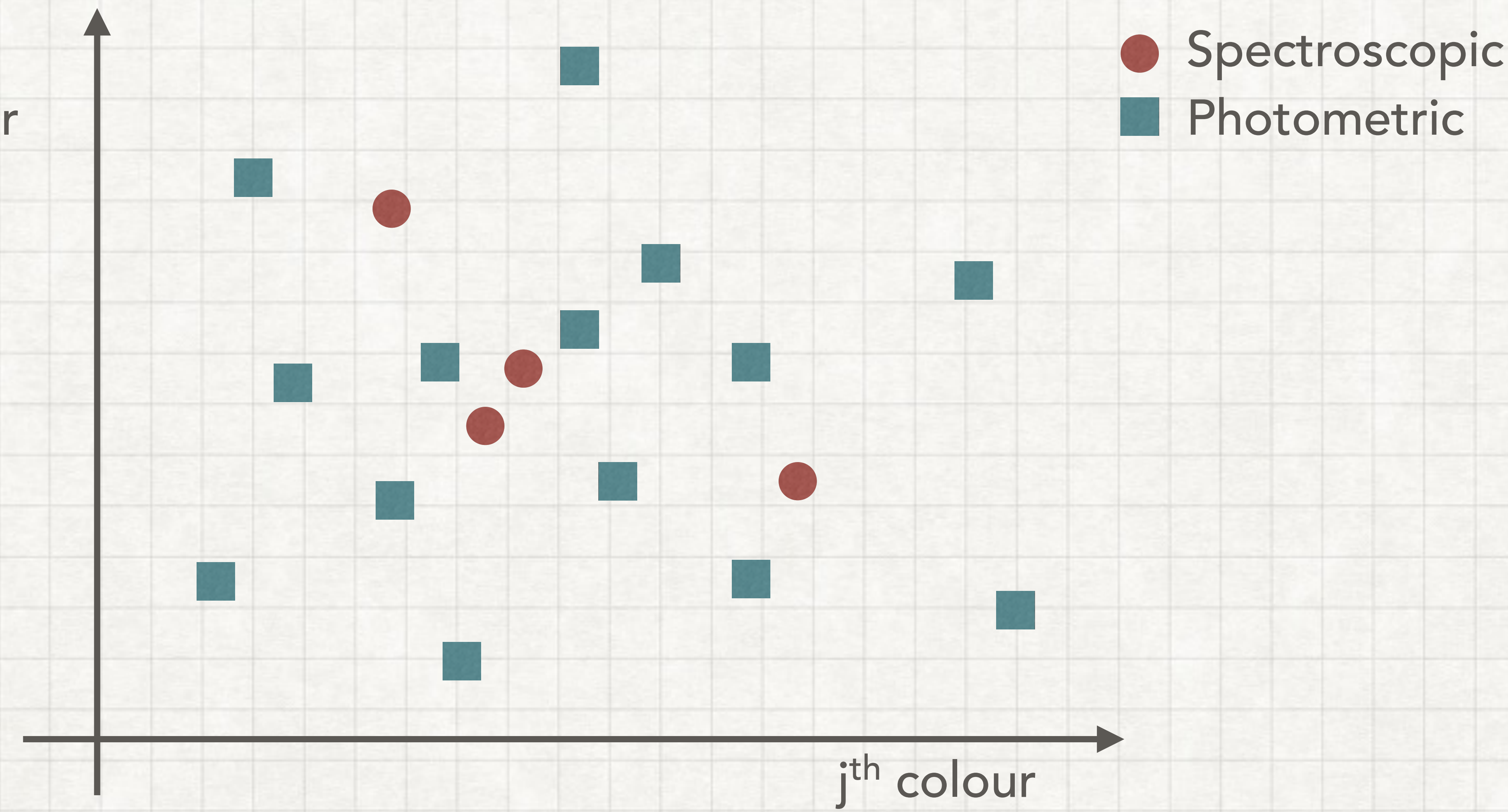
One calibration option is to use (high accuracy) spectroscopic redshifts to construct redshift distributions for the cosmic-shear (i.e. wide-field photometric) galaxies.

↳ Requires a method of mapping spectra onto wide-field galaxies

↳ Requires a representative spec-z sample *or* a method for determining which photometric sources are matched

KIDS DIRECT REDSHIFT CALIBRATION (DIR)

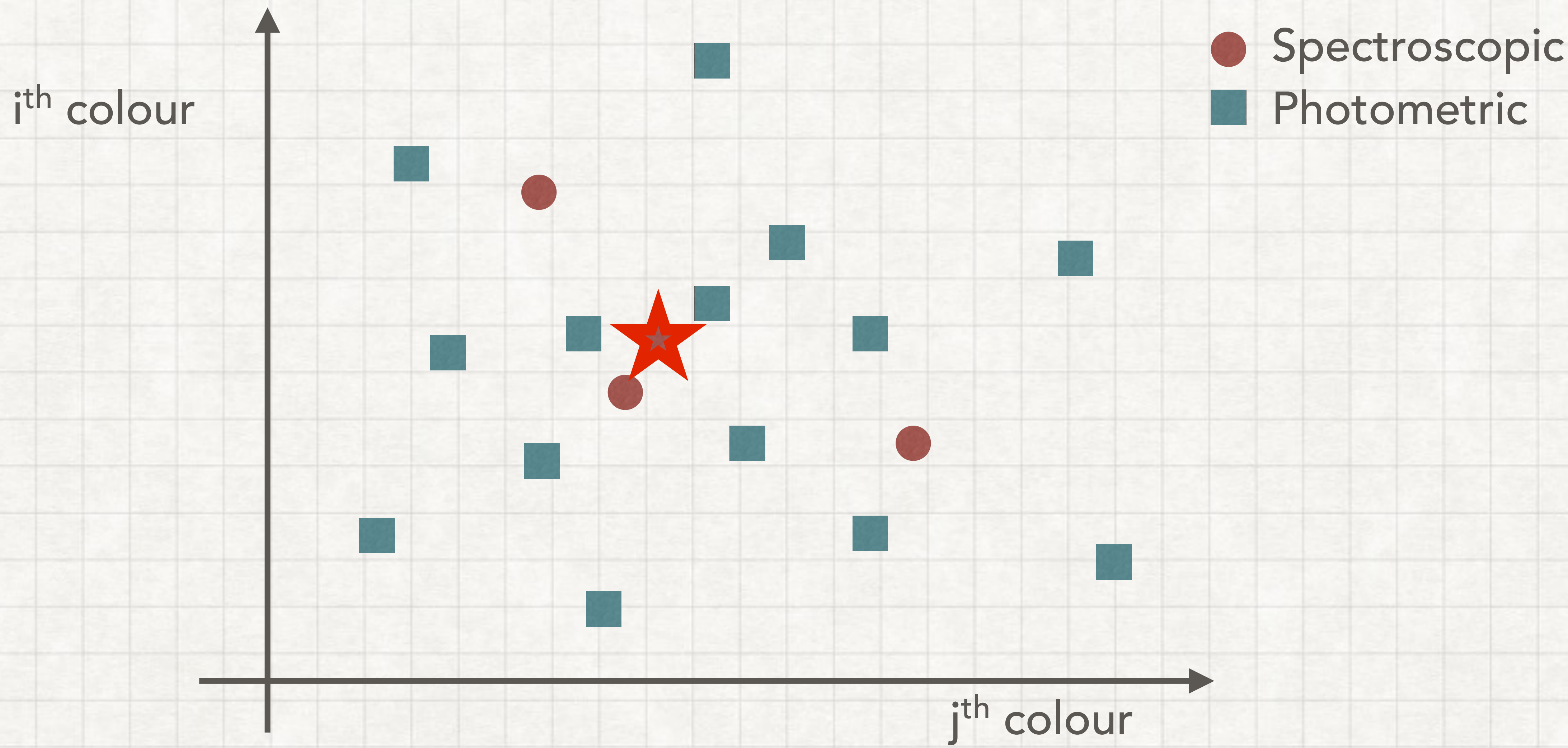
K-NEAREST-NEIGHBOUR ASSOCIATION



The original implementation of direct redshift calibration used kNN to associate spectroscopic and photometric galaxies (Lima et al 2008, Hildebrandt et al 2017,2018)

KIDS DIRECT REDSHIFT CALIBRATION (DIR)

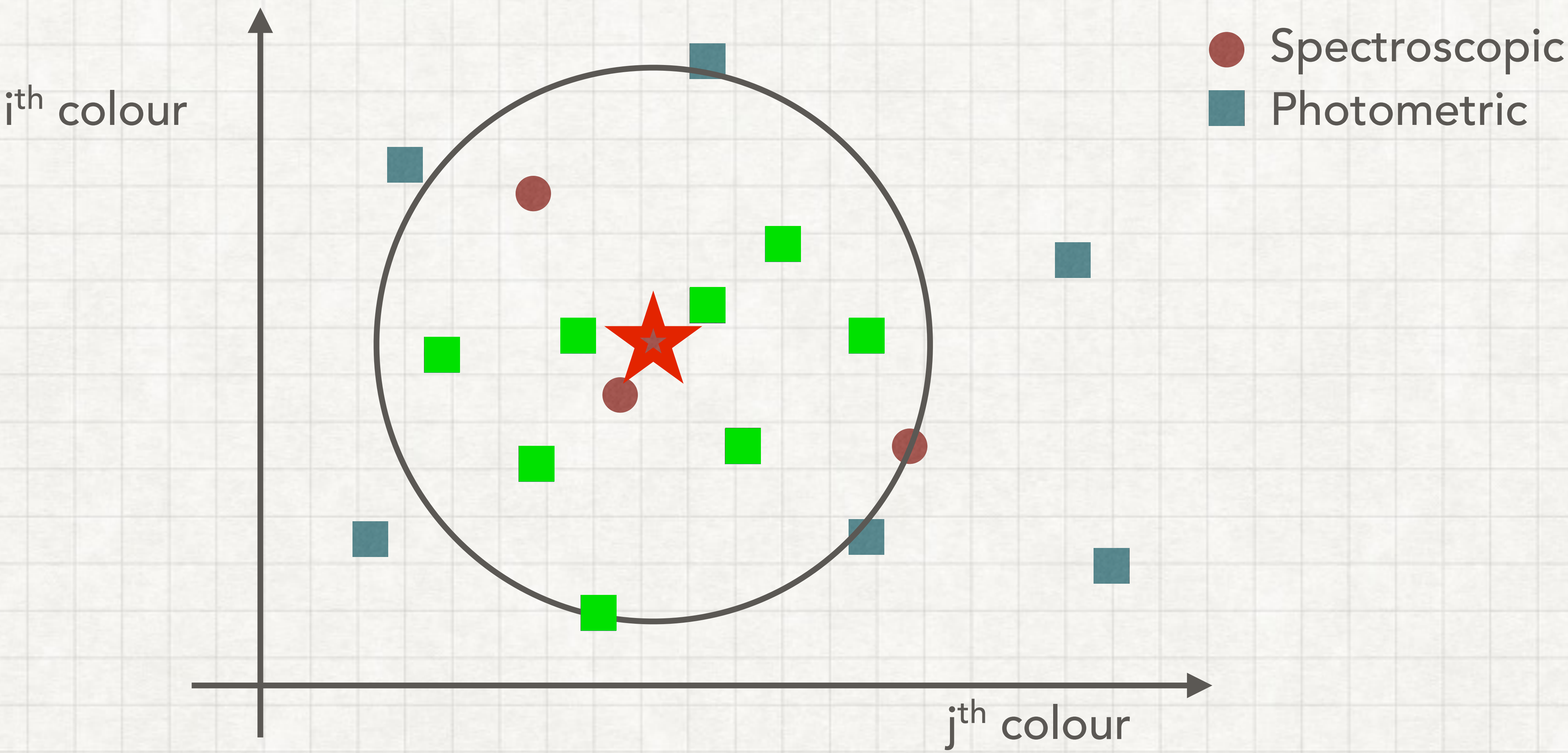
K-NEAREST-NEIGHBOUR ASSOCIATION



The original implementation of direct redshift calibration used kNN to associate spectroscopic and photometric galaxies (Lima et al 2008, Hildebrandt et al 2017,2018)

KIDS DIRECT REDSHIFT CALIBRATION (DIR)

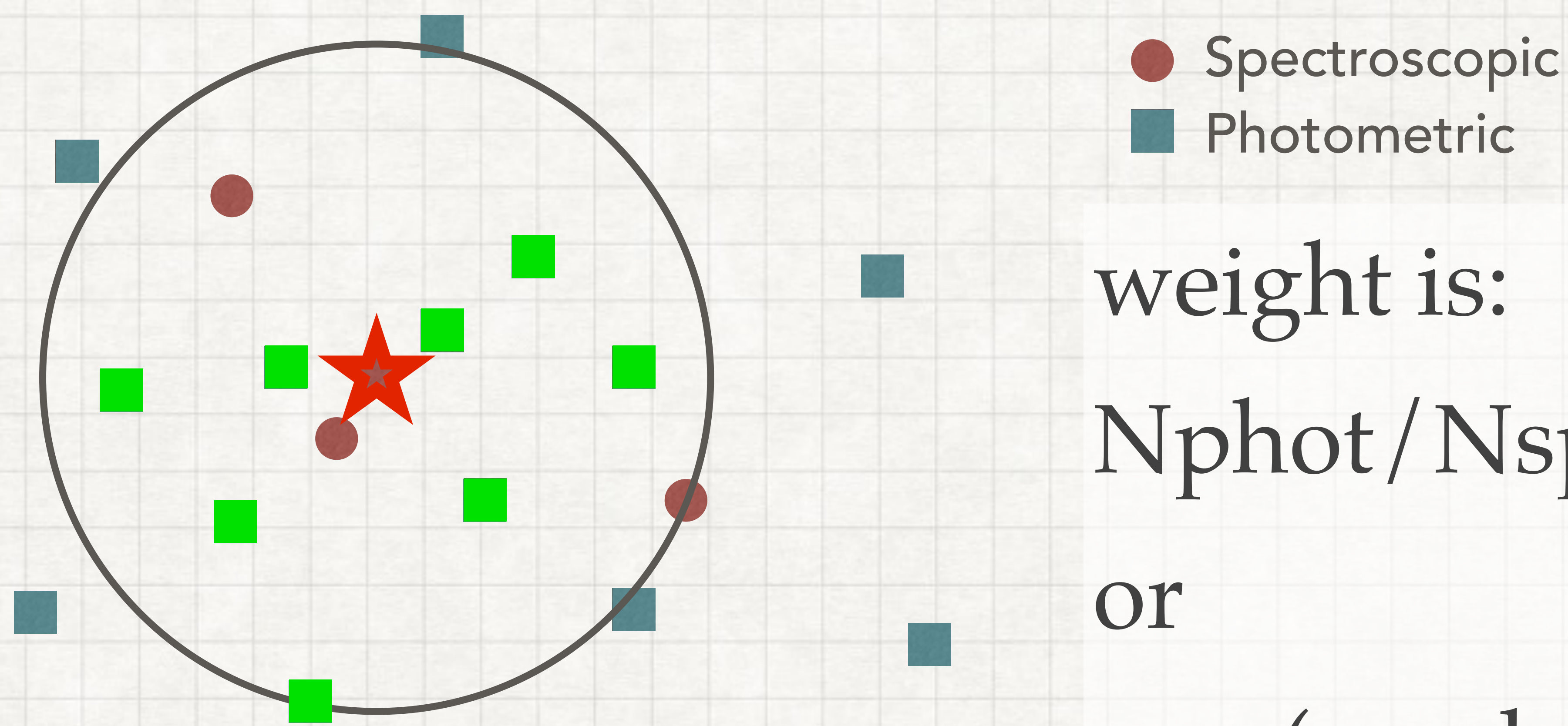
K-NEAREST-NEIGHBOUR ASSOCIATION



The original implementation of direct redshift calibration used kNN to associate spectroscopic and photometric galaxies (Lima et al 2008, Hildebrandt et al 2017,2018)

KIDS DIRECT REDSHIFT CALIBRATION (DIR)

K-NEAREST-NEIGHBOUR ASSOCIATION



weight is:

Nphot/Nspec

or

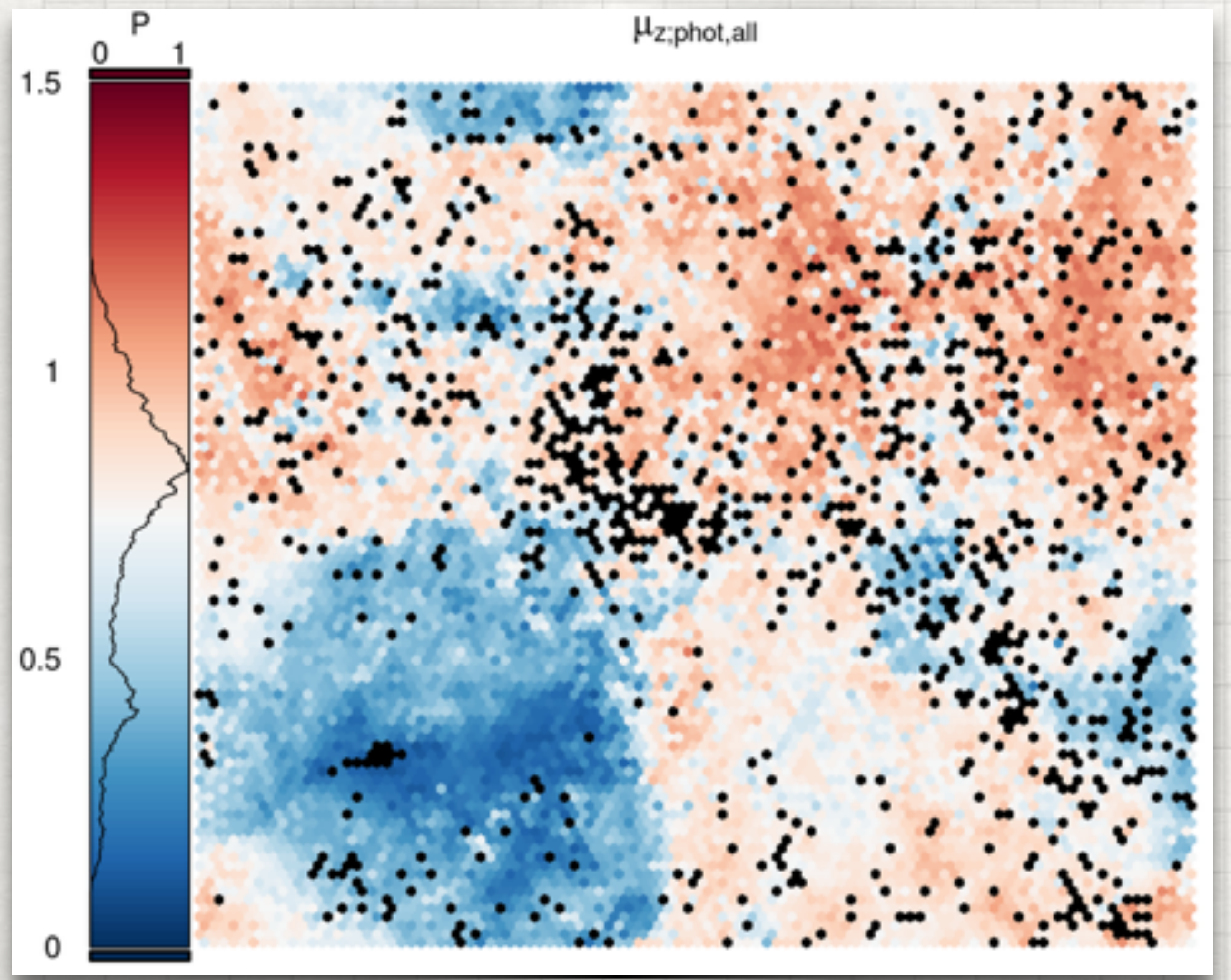
sum(w_phot)/Nspec

The original implementation of direct redshift calibration used kNN to associate spectroscopic and photometric galaxies (Lima et al 2008, Hildebrandt et al 2017,2018)

KIDS SOM REDSHIFT CALIBRATION (SOM)

A NOVEL APPROACH

Self organising maps present another method of associating galaxies in distinct catalogues. The weighting mathematics remains the same, but the form of the association changes.



Wright et al (submitted)

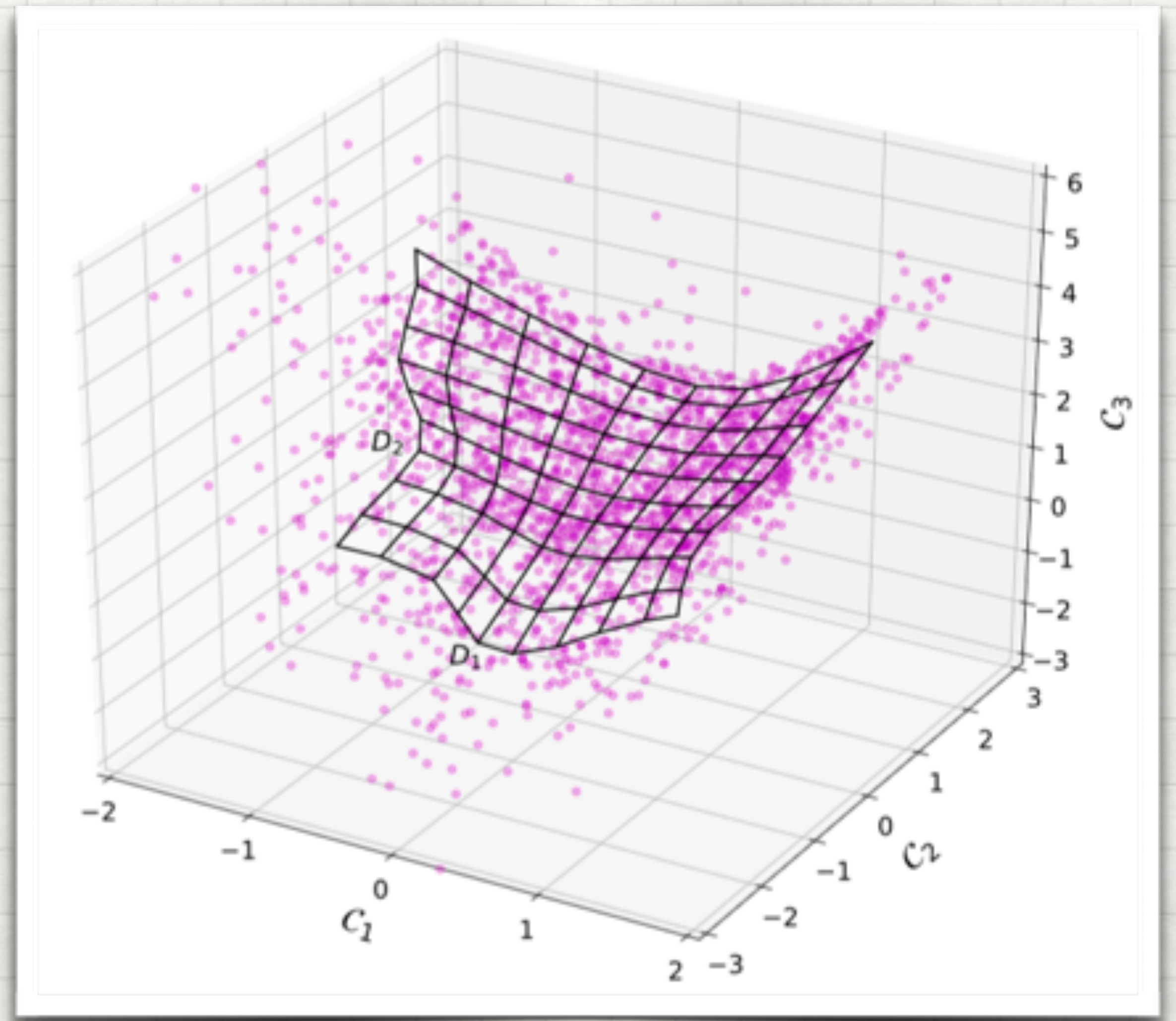
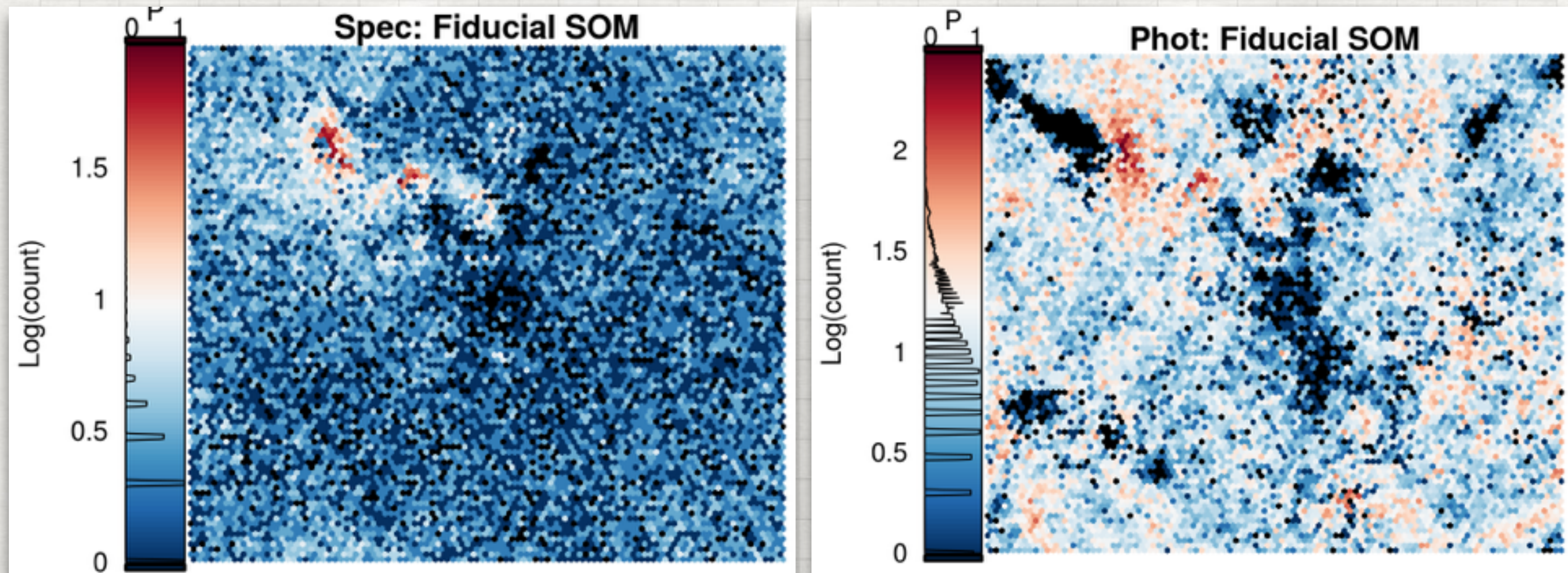


Image credit: Davidzon et al (2019)

REDSHIFT CALIBRATION WITH SELF-ORGANISING MAPS

A NOVEL APPROACH

Associations are based on the self-similarity between the sources in respect to the colour-colour manifold. Allows variable Nspec. Number of neighbours choice is converted to SOM pixel choice.



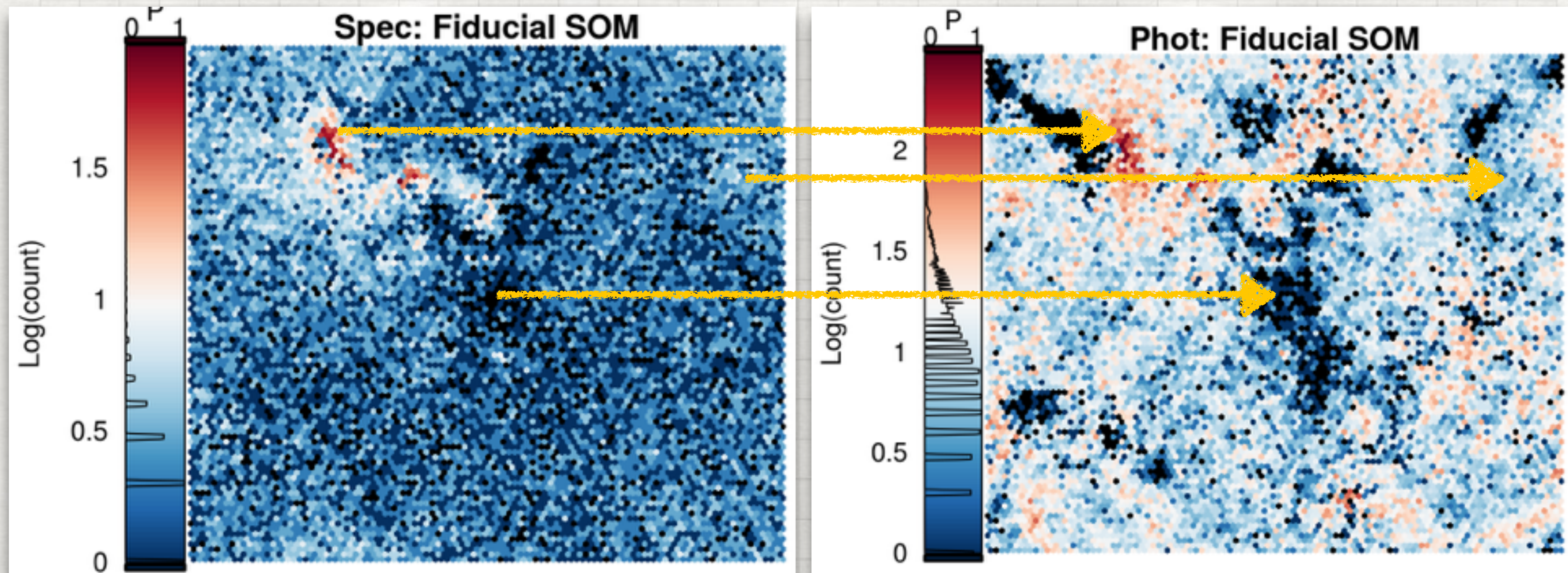
Direct knowledge of cells with shear sources and no spec-z

Wright et al (submitted)

REDSHIFT CALIBRATION WITH SELF-ORGANISING MAPS

A NOVEL APPROACH

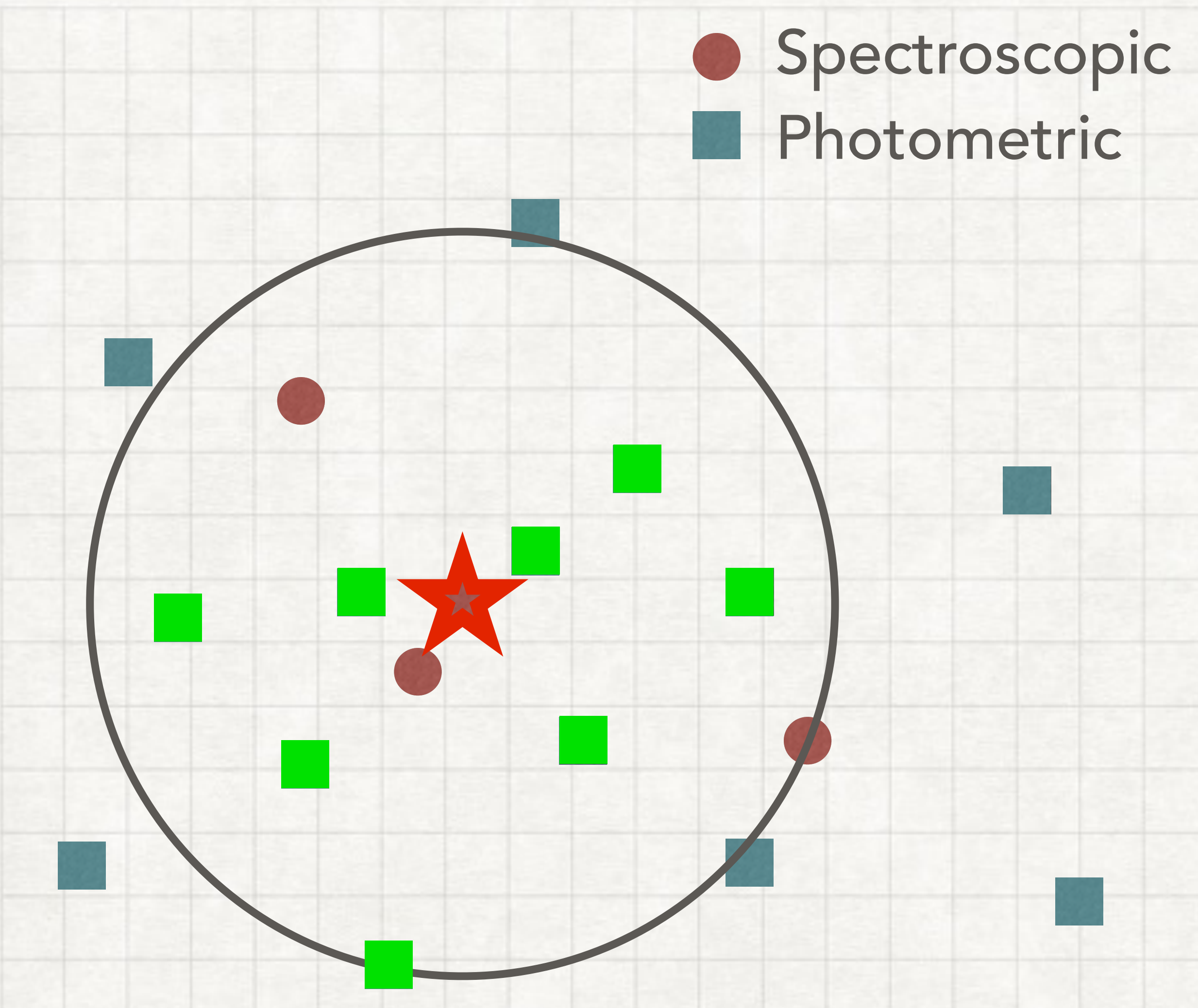
Associations are based on the self-similarity between the sources in respect to the colour-colour manifold. Allows variable Nspec. Number of neighbours choice is converted to SOM pixel choice.

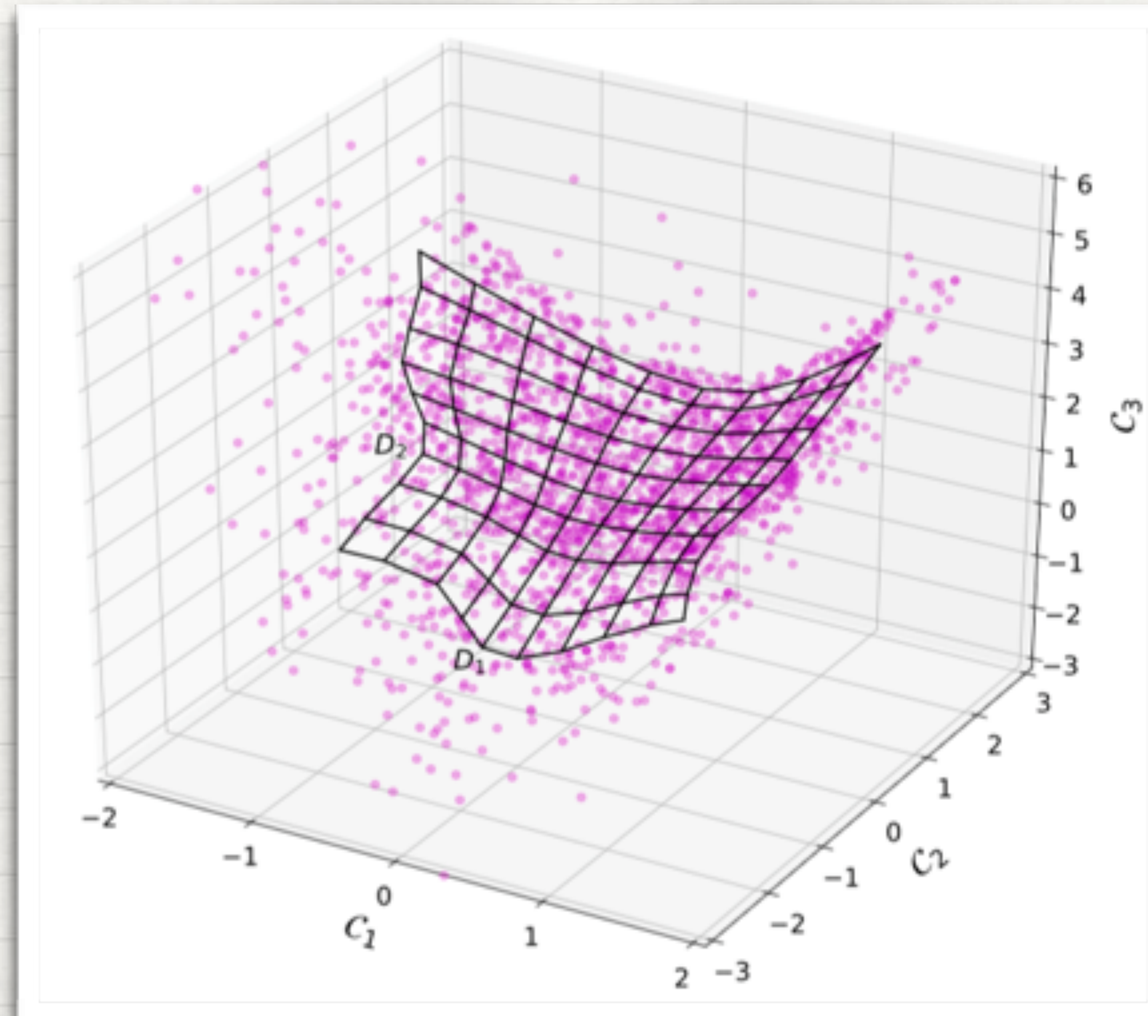


Direct knowledge of cells with shear sources and no spec-z

Wright et al (submitted)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



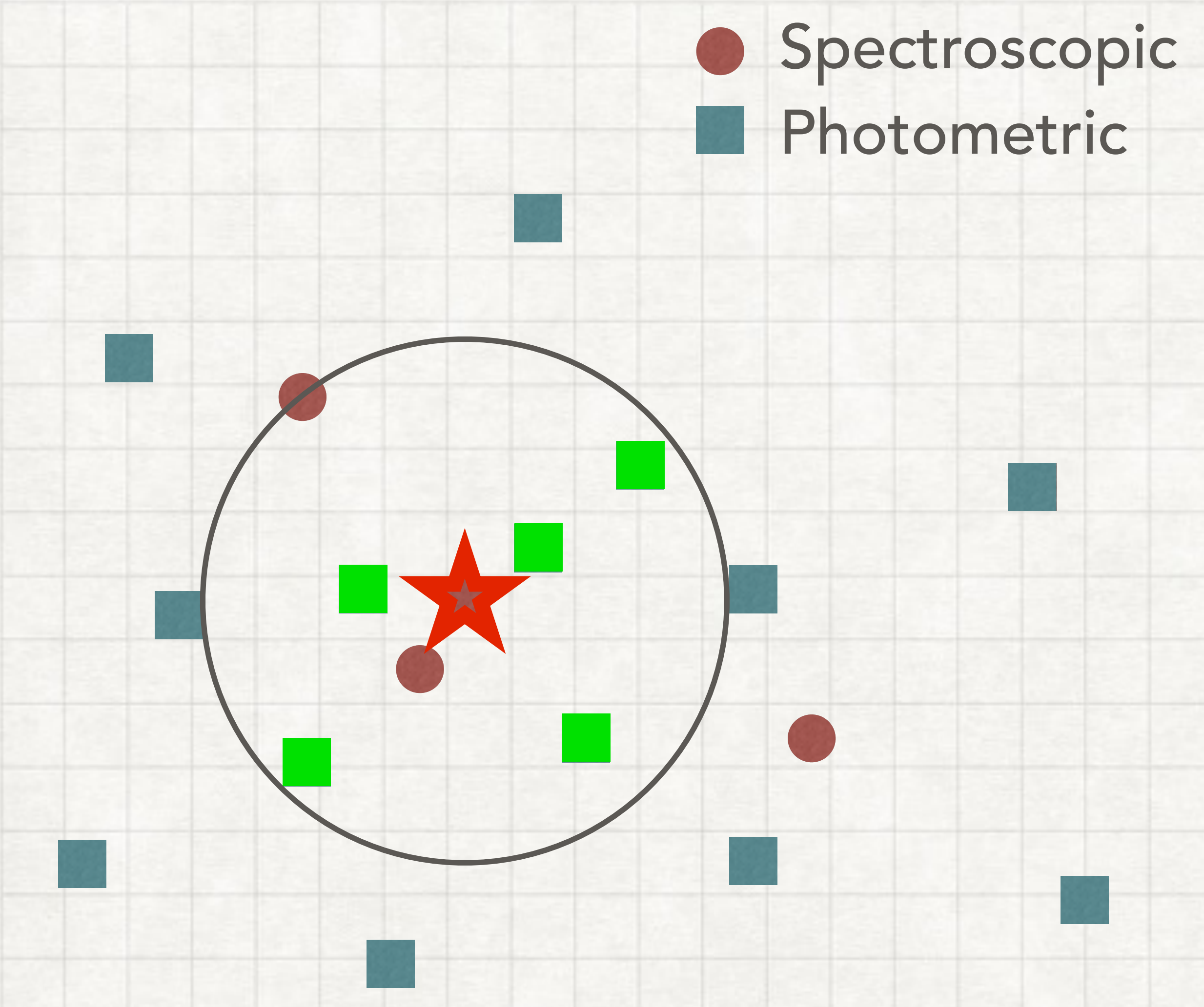


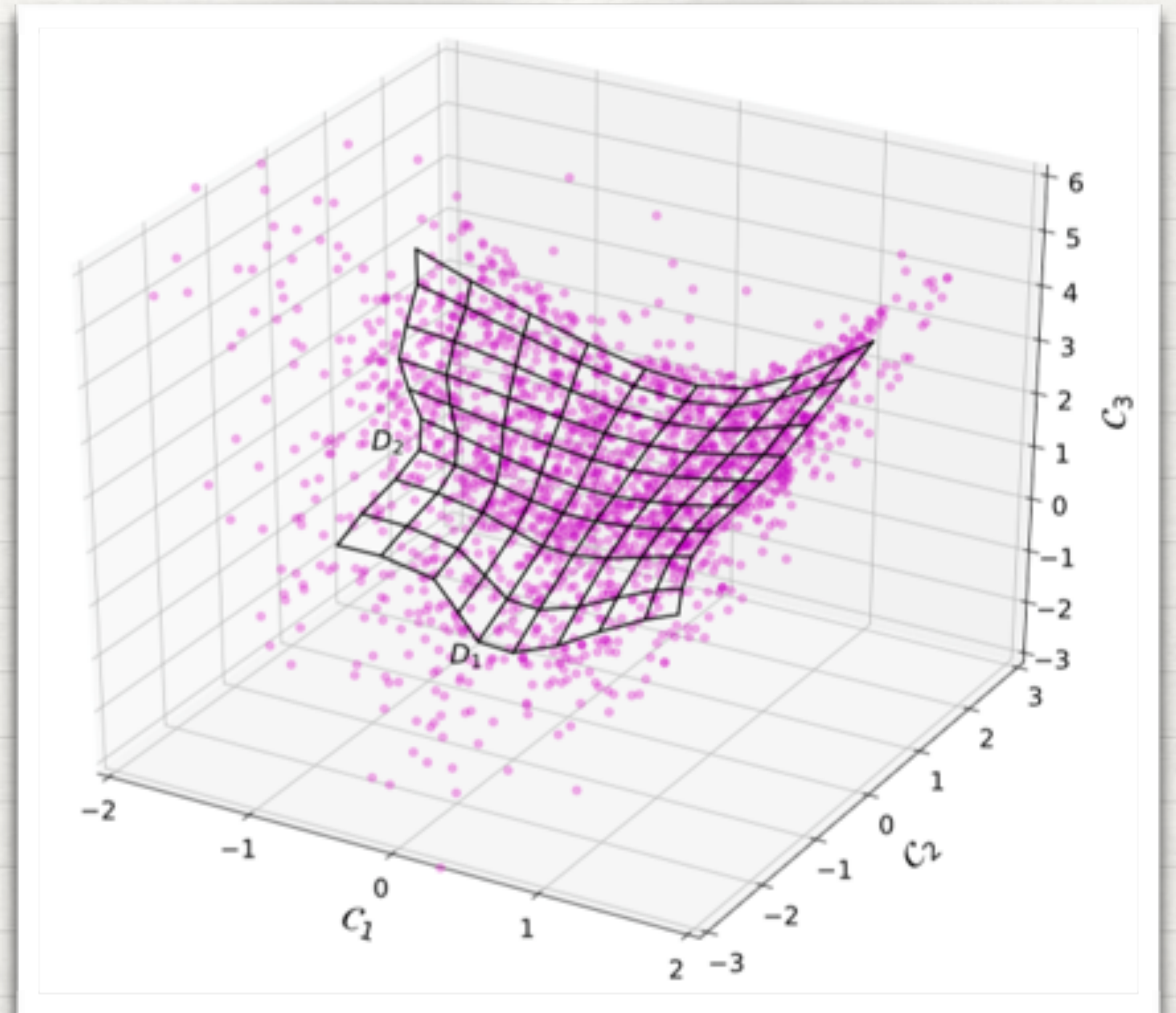
The choice of pixel scale imposes limitation on the mapping of the hyper-surface.

↳ Clustering the pixel data avoids this while not biasing the calibration

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



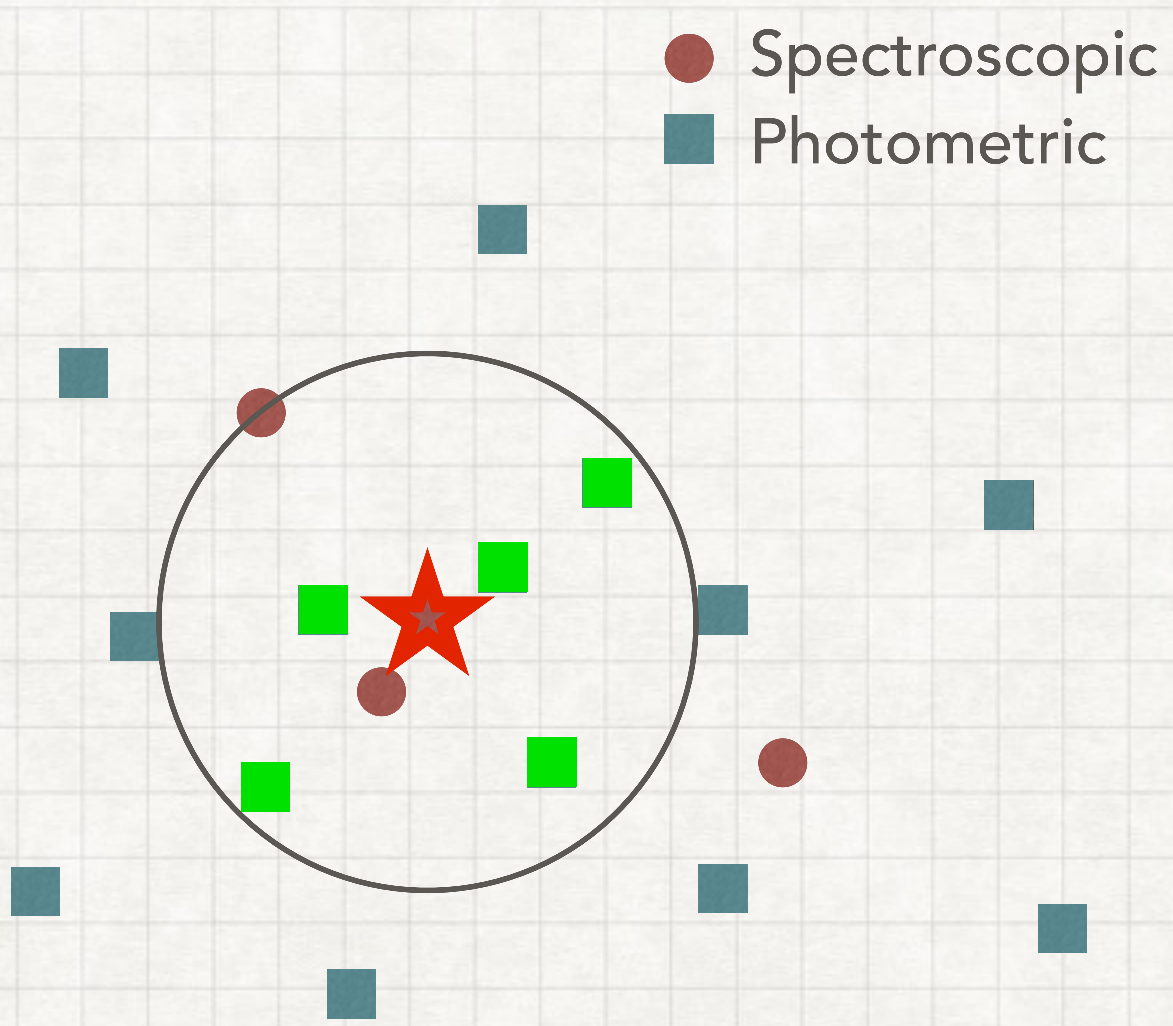


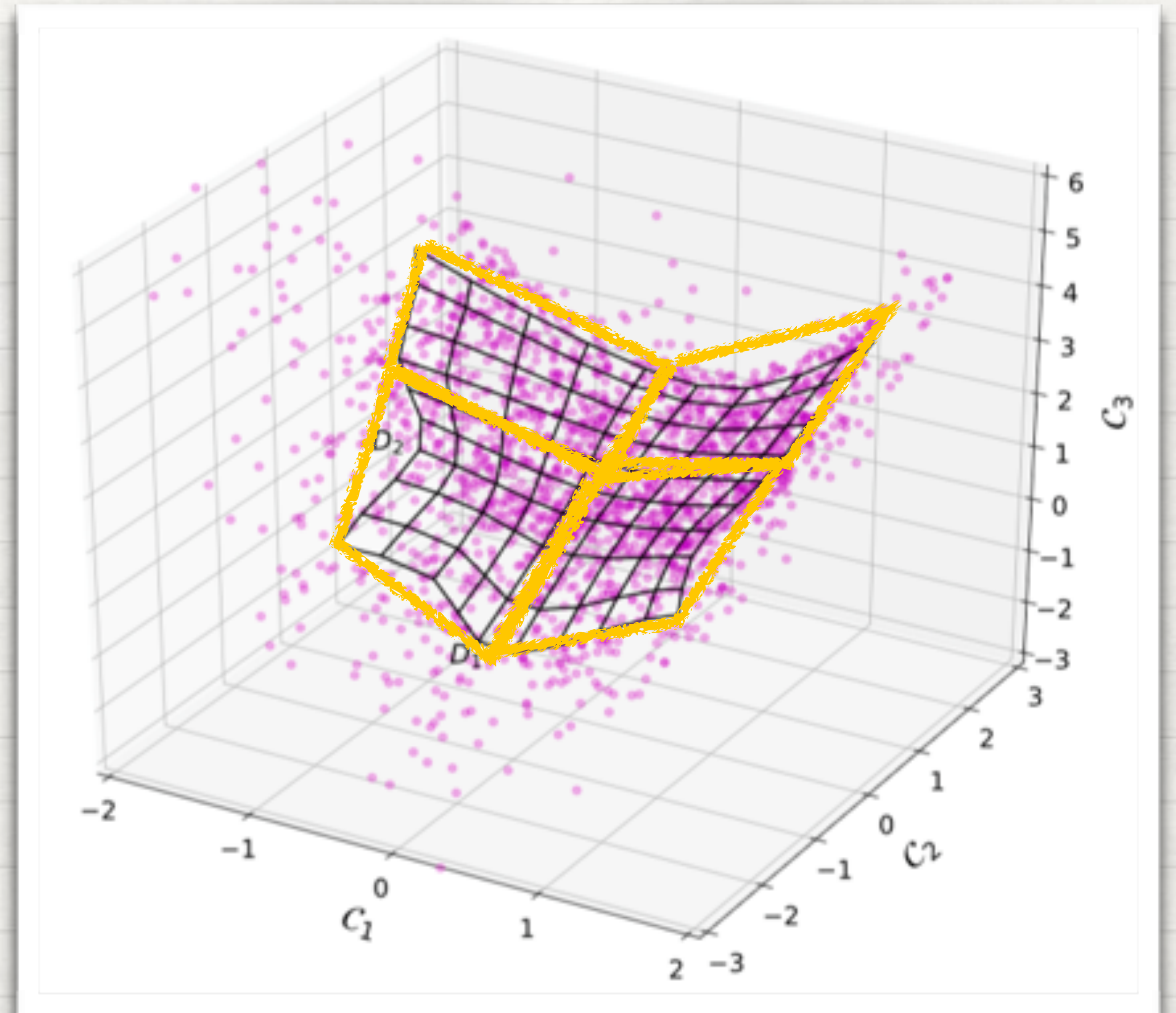
The choice of pixel scale imposes limitation on the mapping of the hyper-surface.

↳ Clustering the pixel data avoids this while not biasing the calibration

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



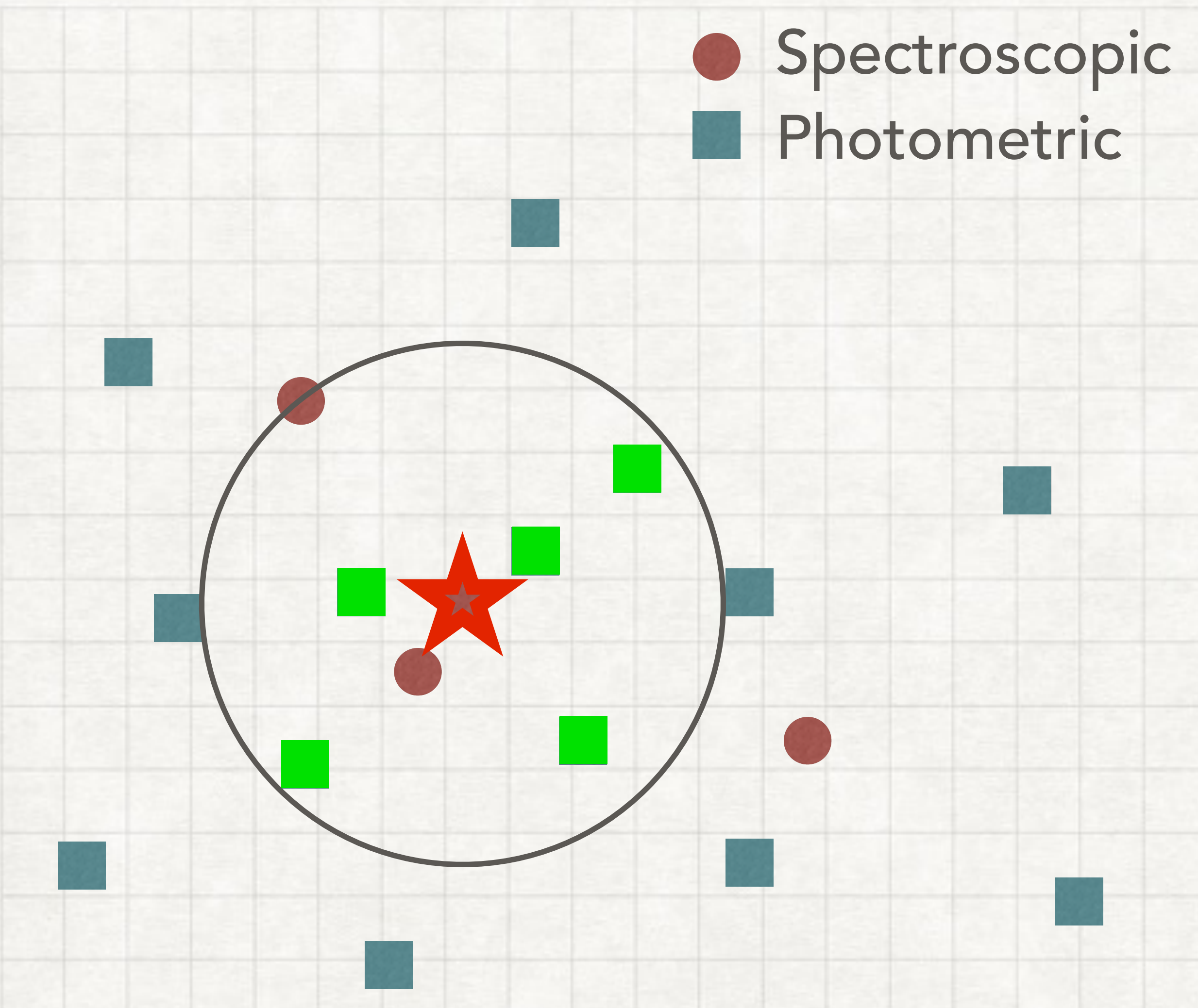


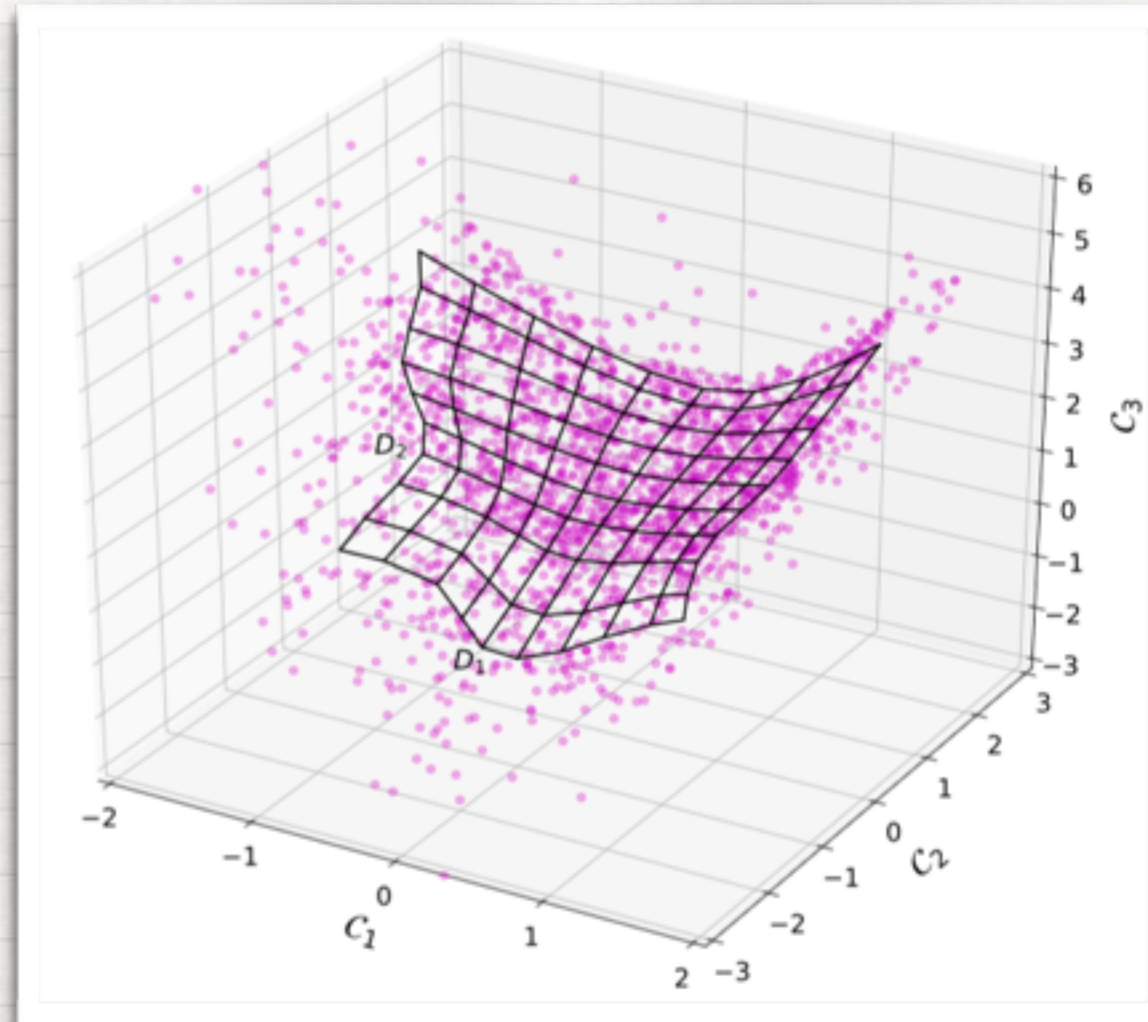
The choice of pixel scale imposes limitation on the mapping of the hyper-surface.

↳ Clustering the pixel data avoids this while not biasing the calibration

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



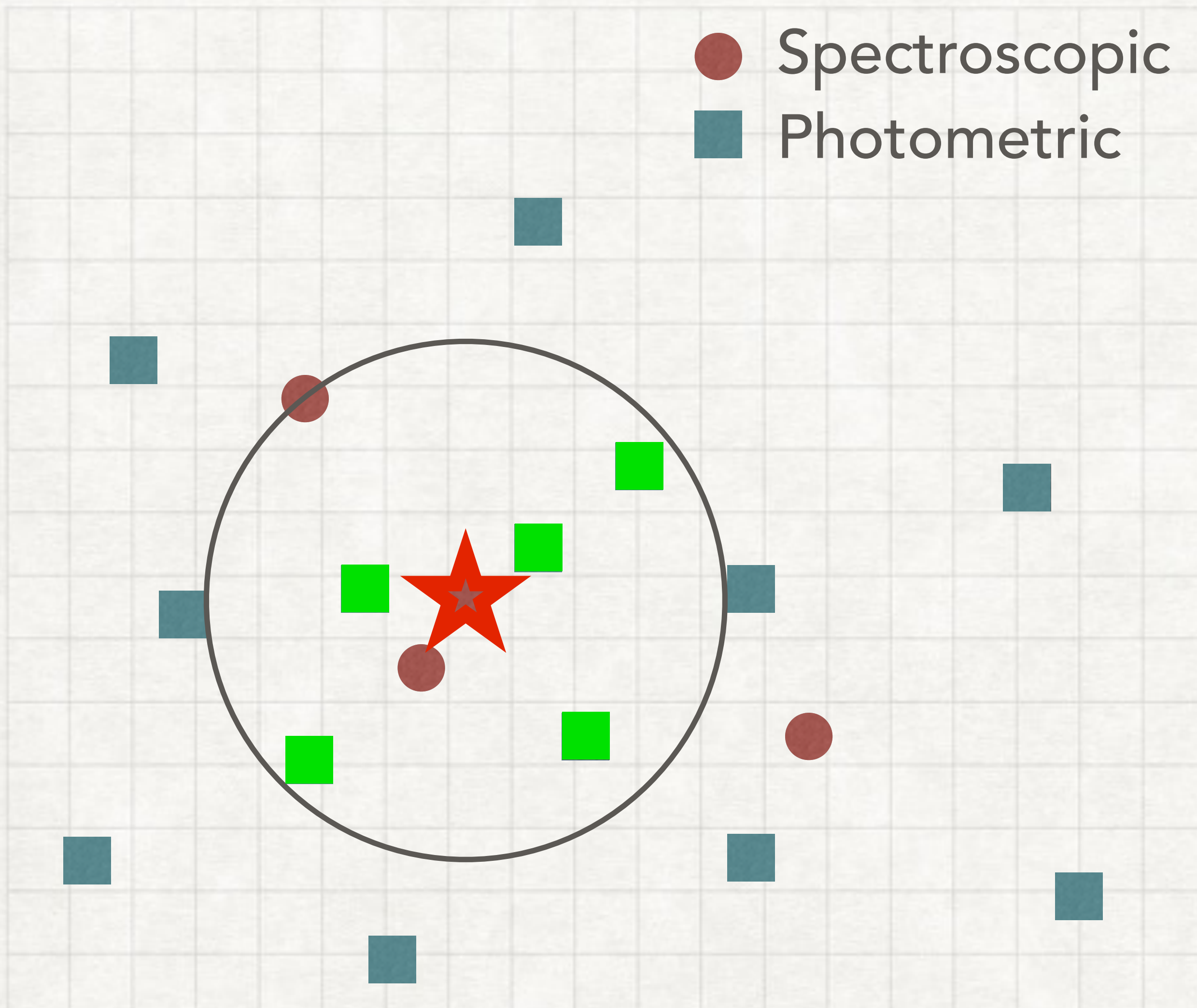


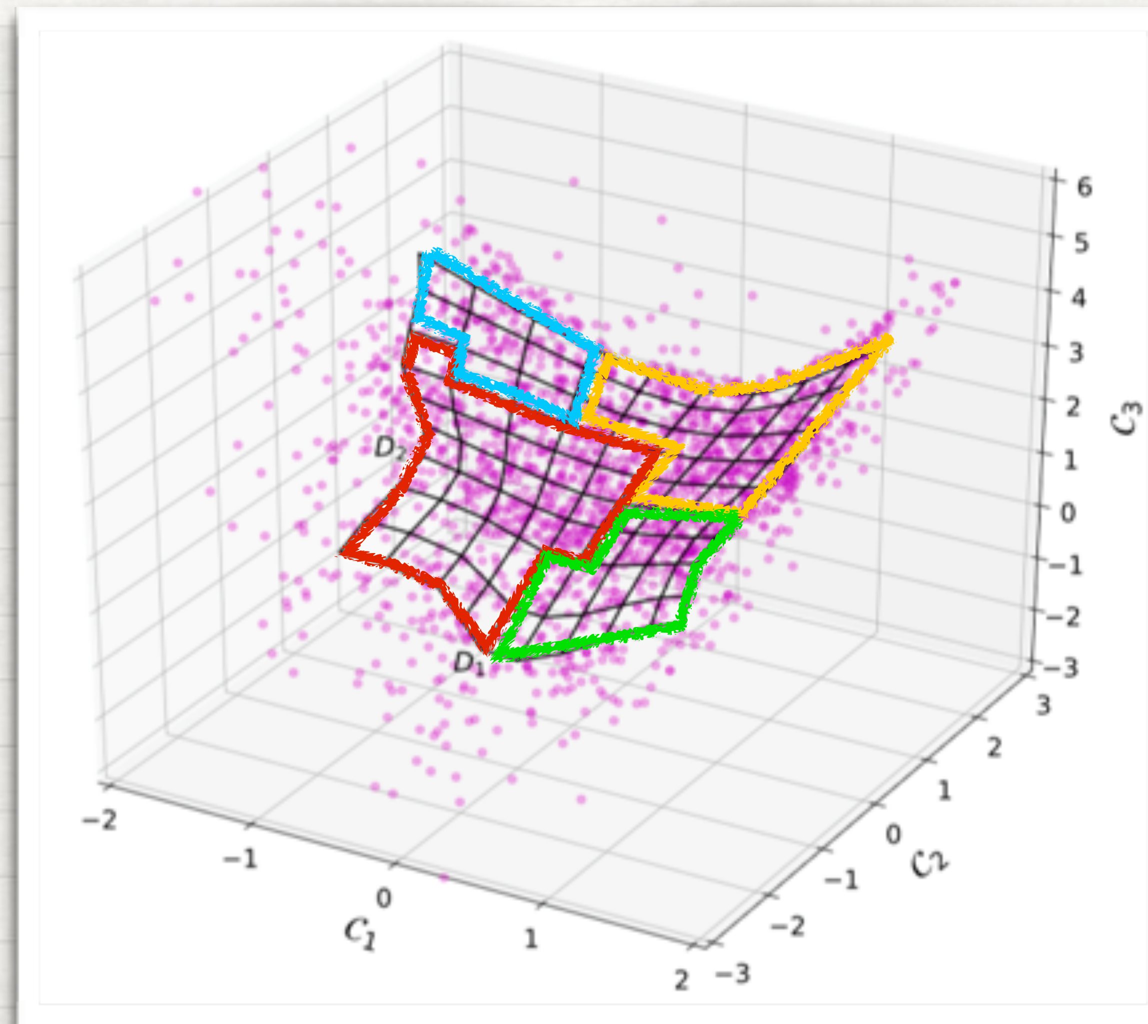
The choice of pixel scale imposes limitation on the mapping of the hyper-surface.

↳ Clustering the pixel data avoids this while not biasing the calibration

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



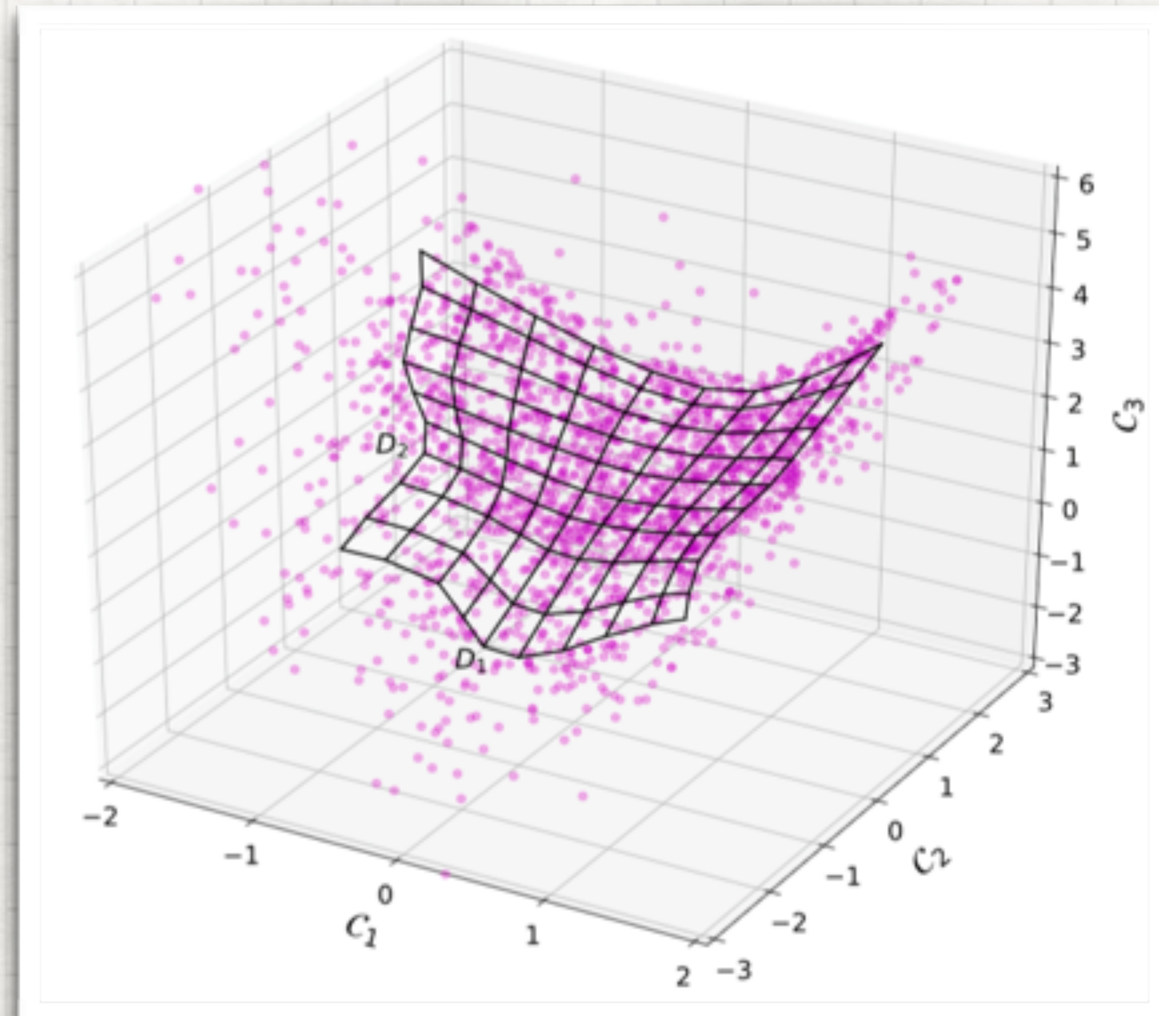


The choice of pixel scale imposes limitation on the mapping of the hyper-surface.

↳ Clustering the pixel data avoids this while not biasing the calibration

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING

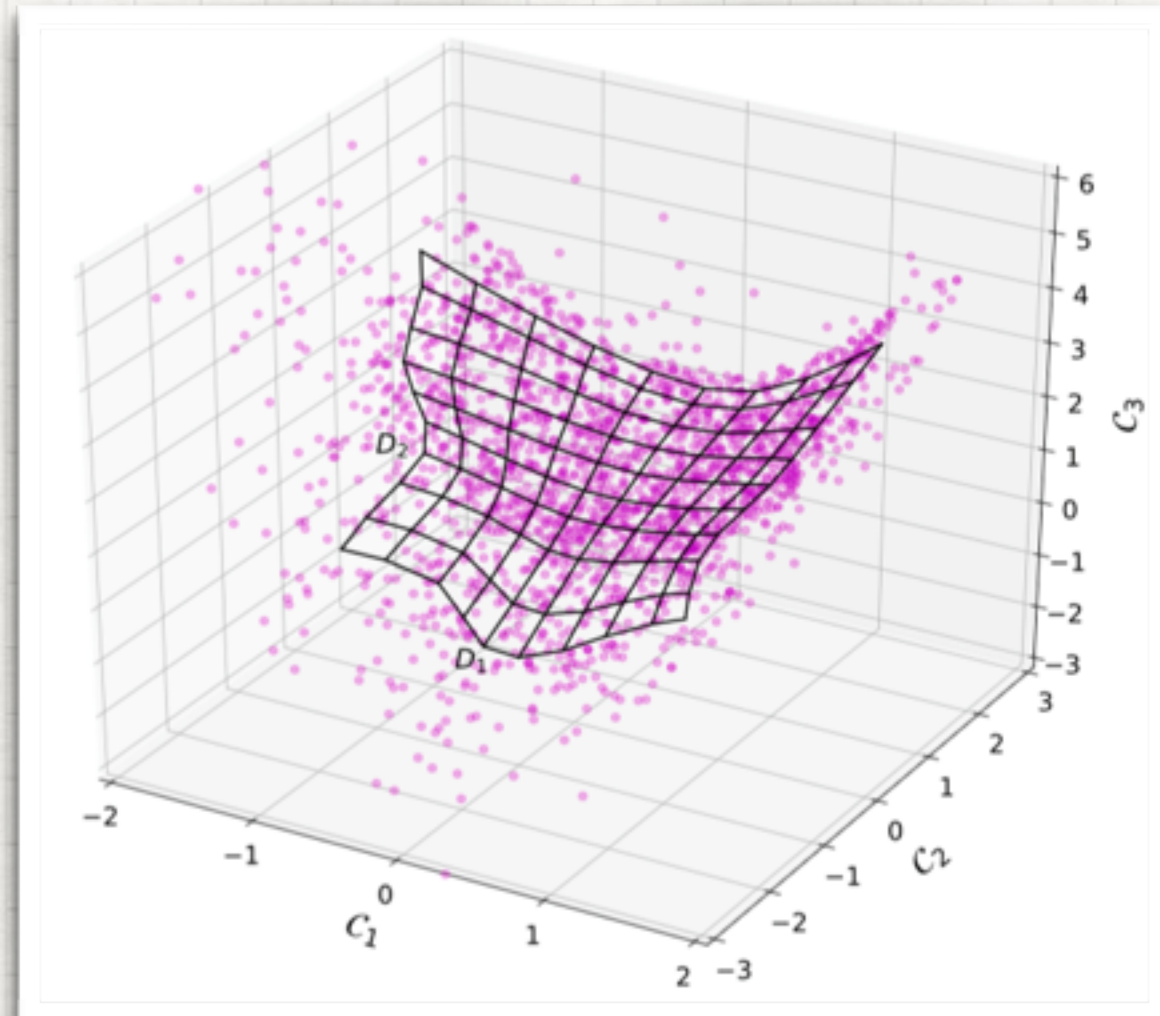


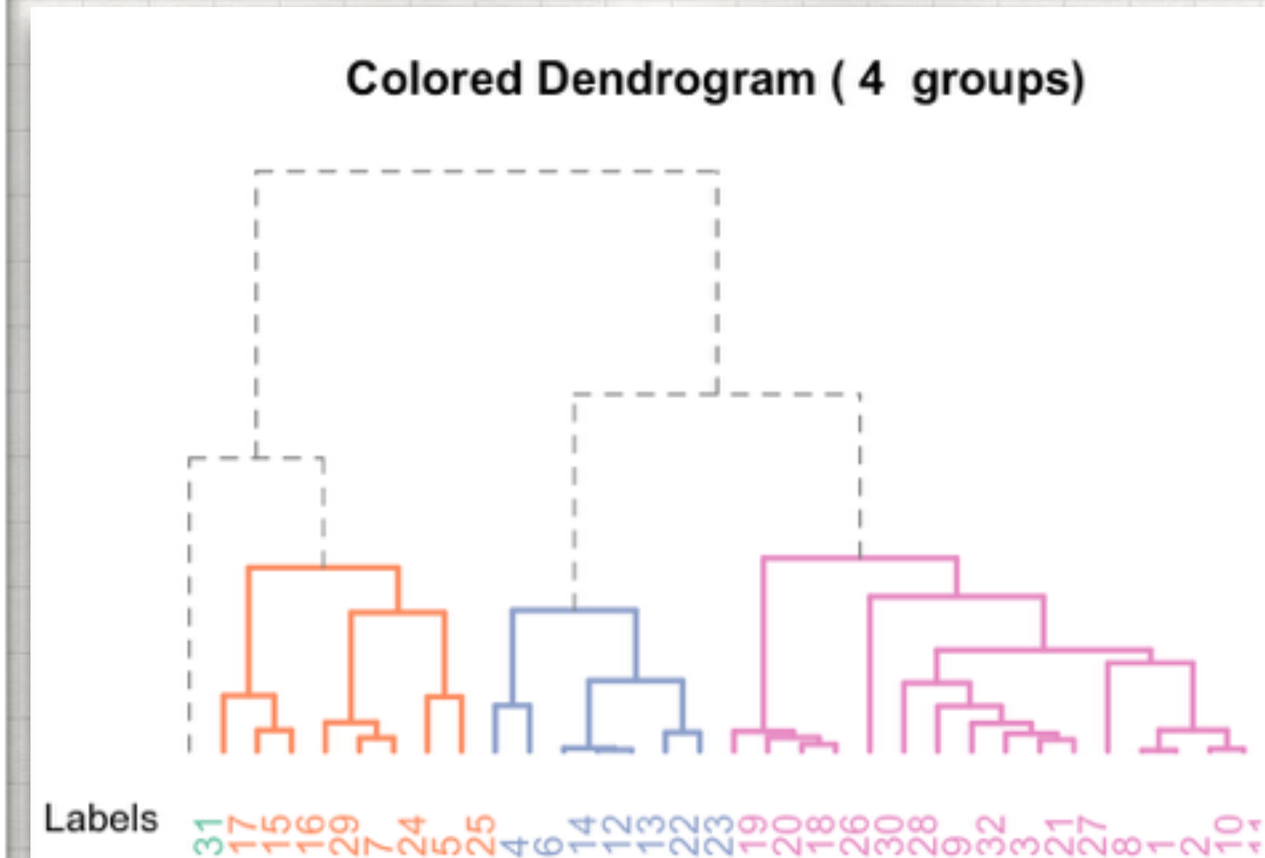
The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



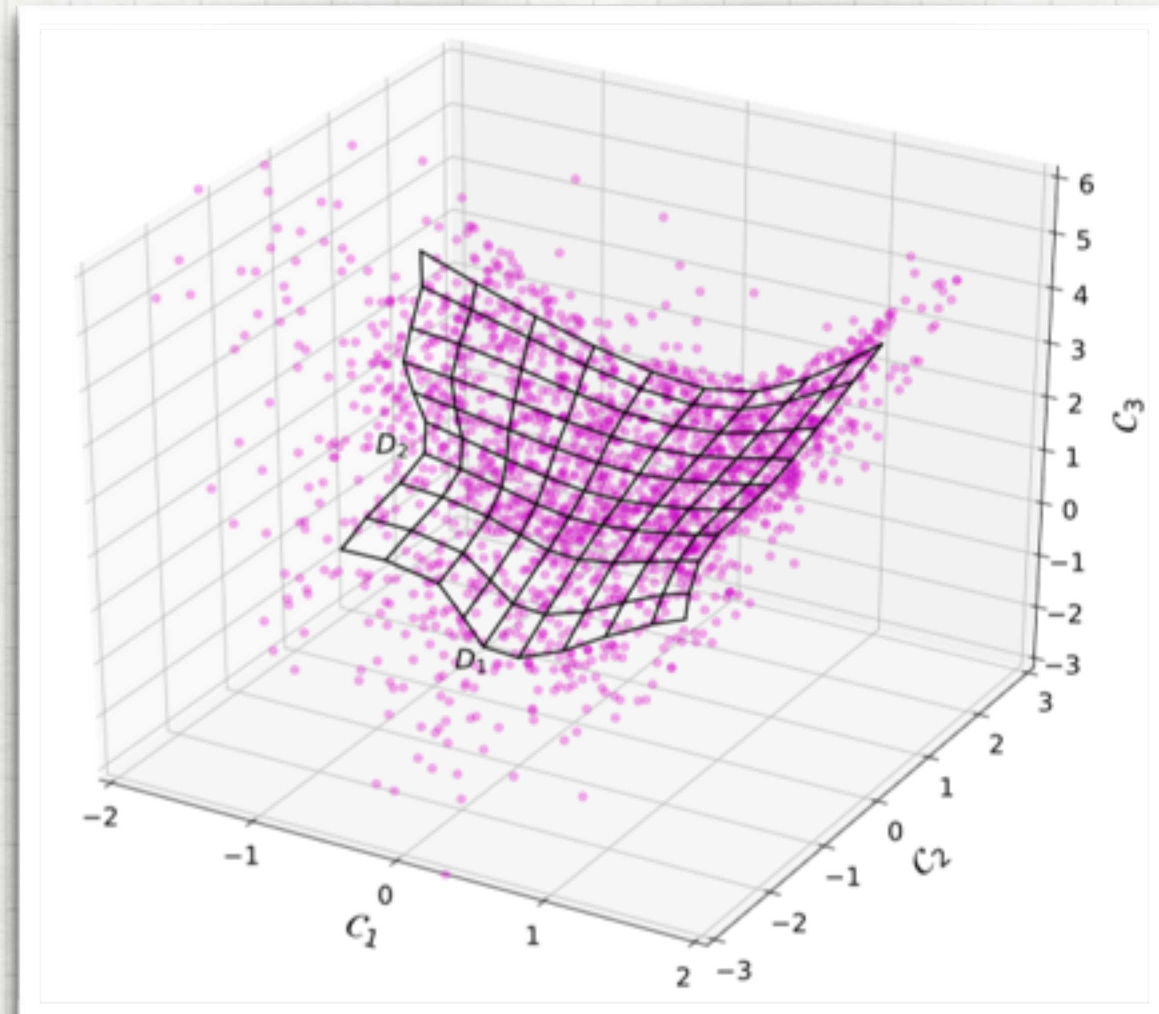


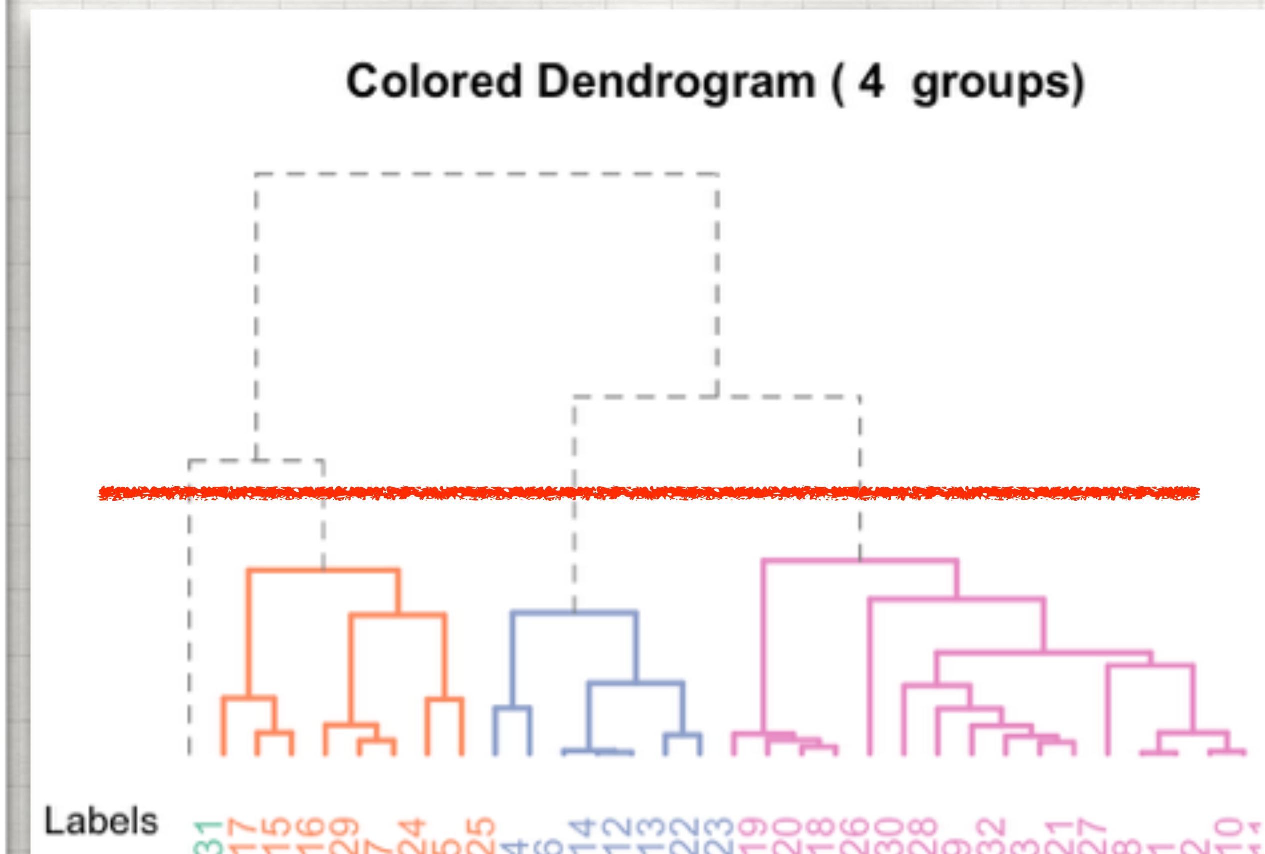
The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



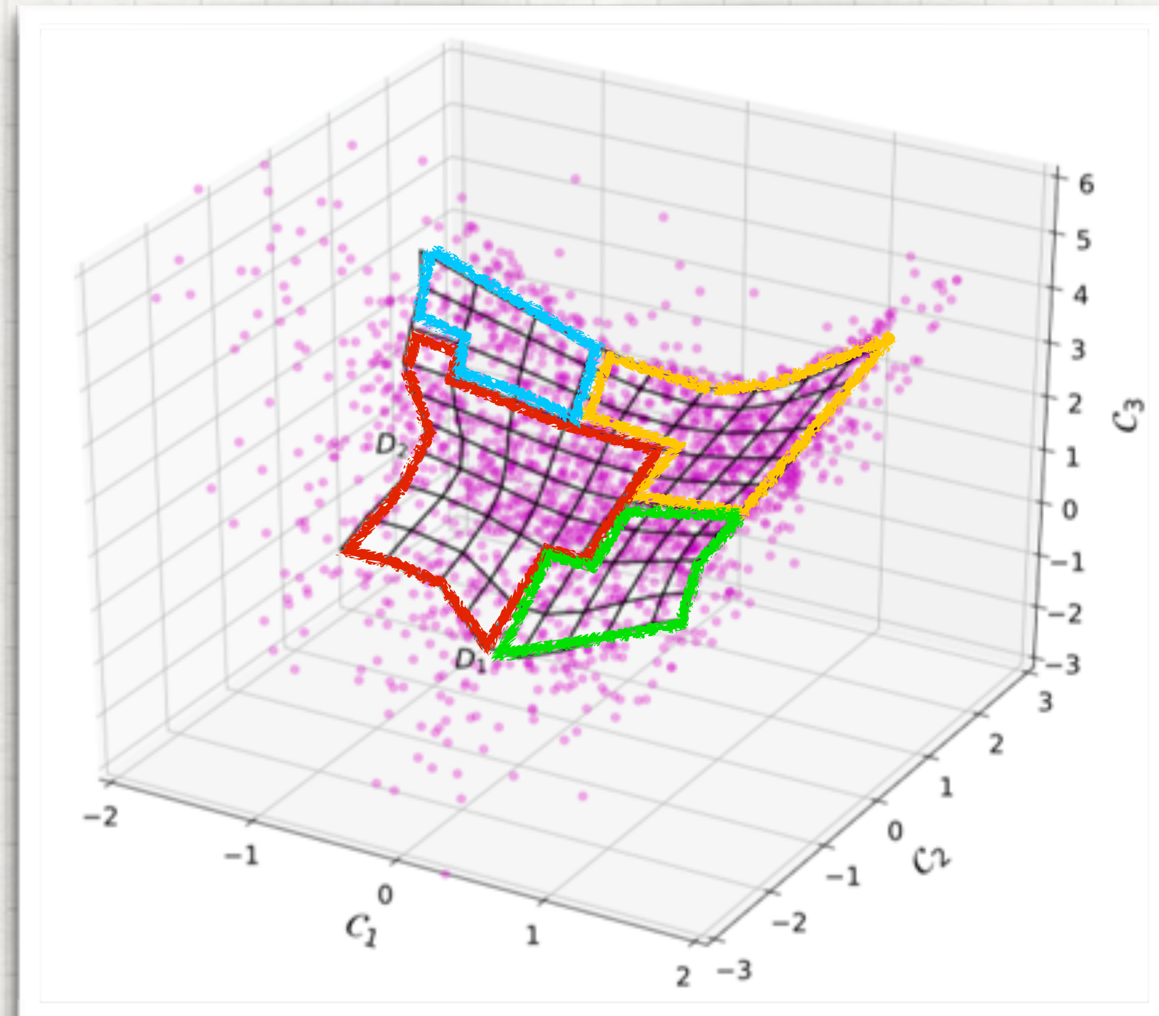


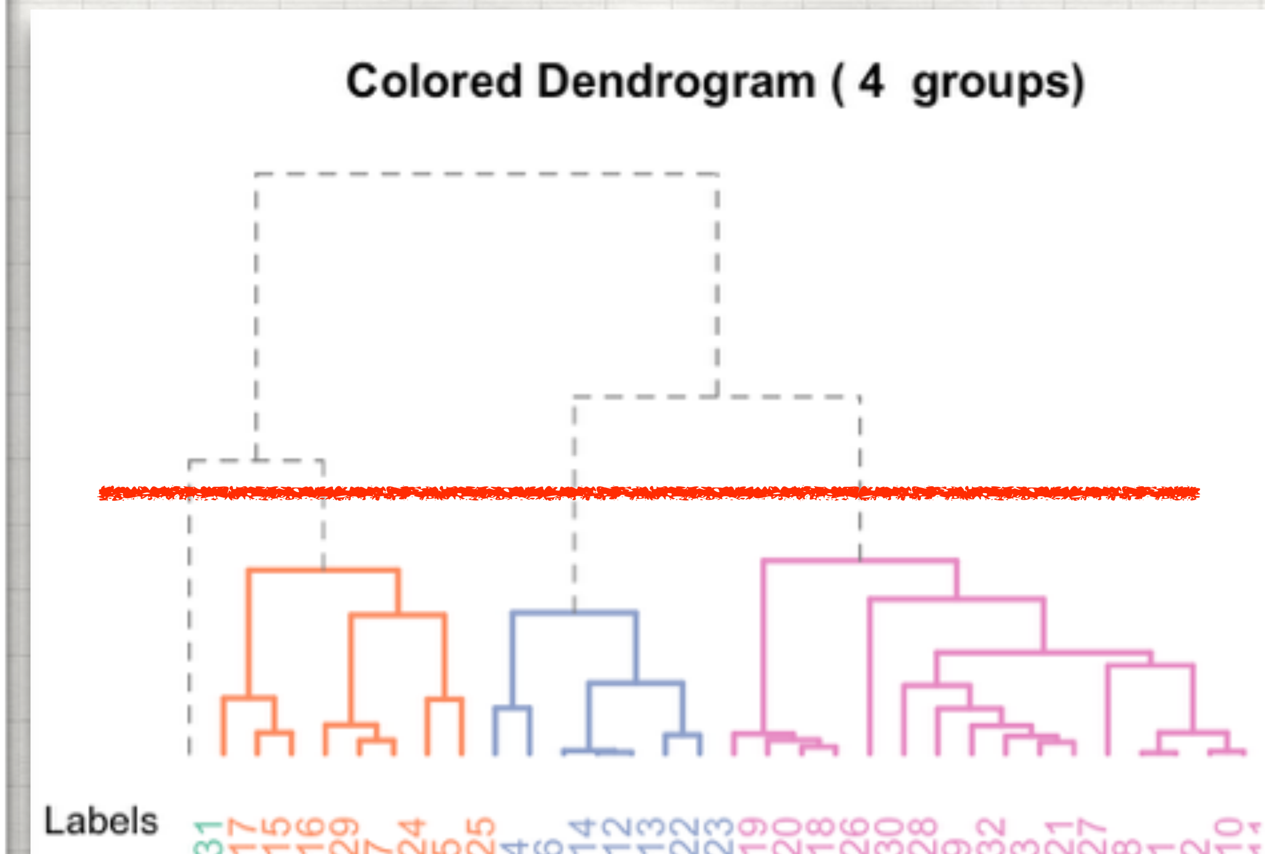
The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



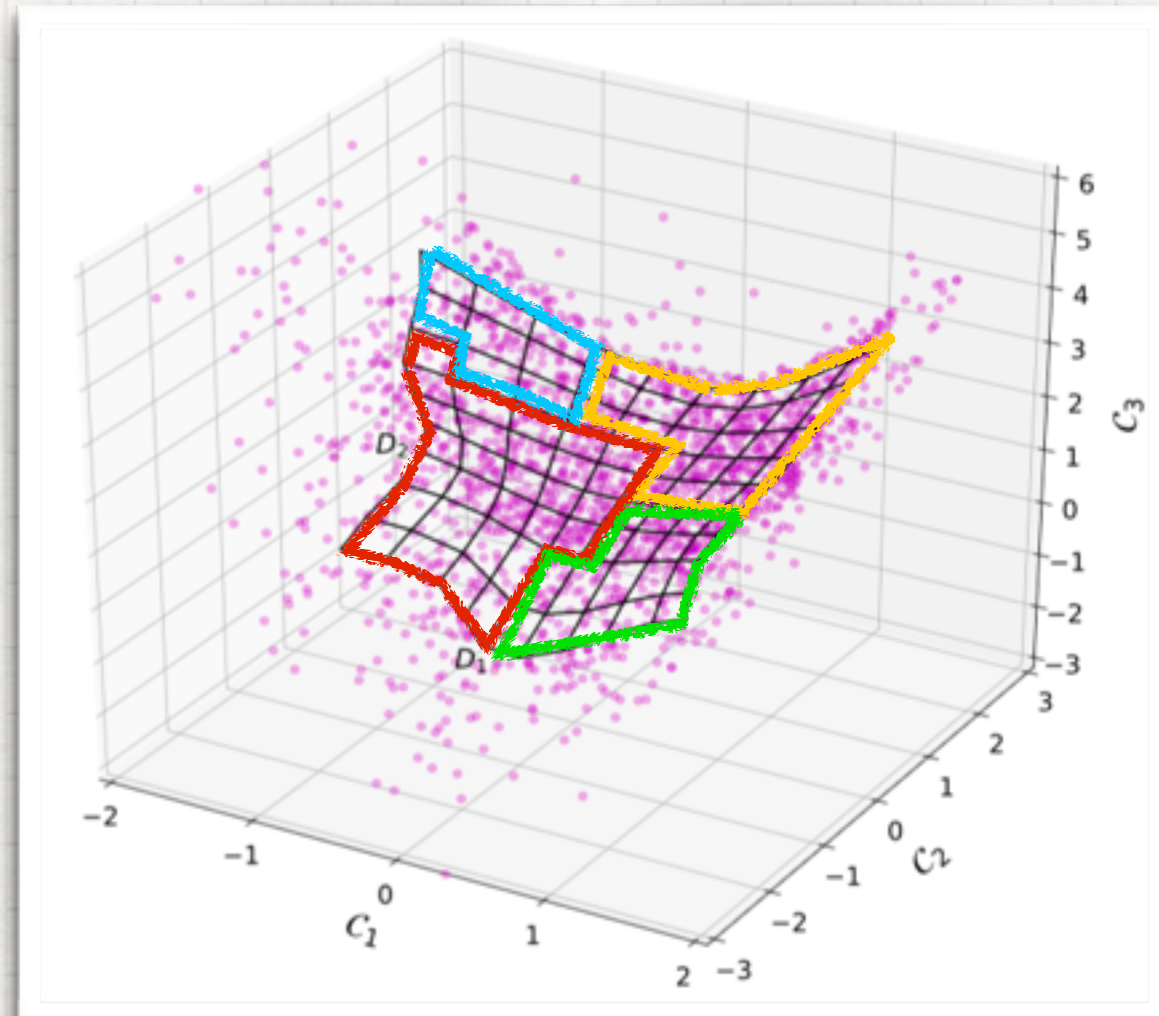


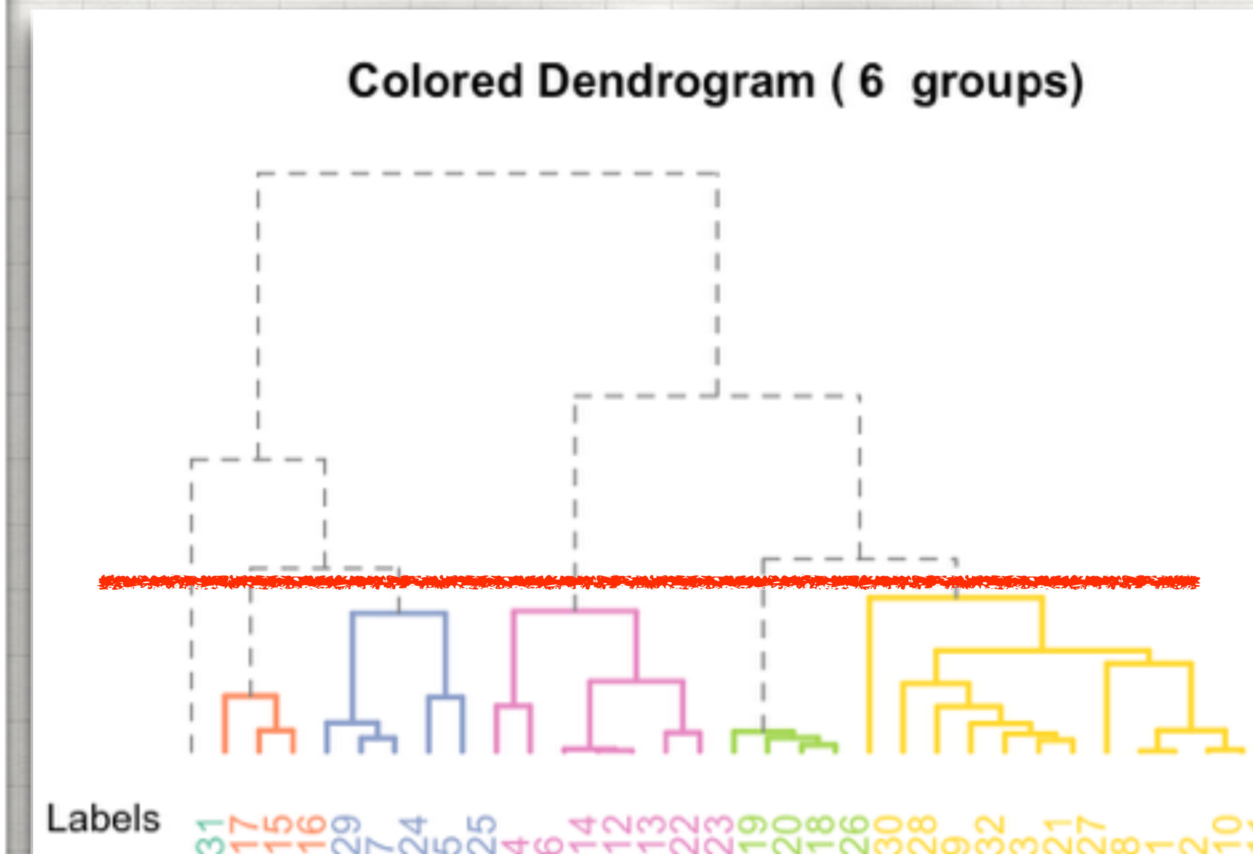
The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



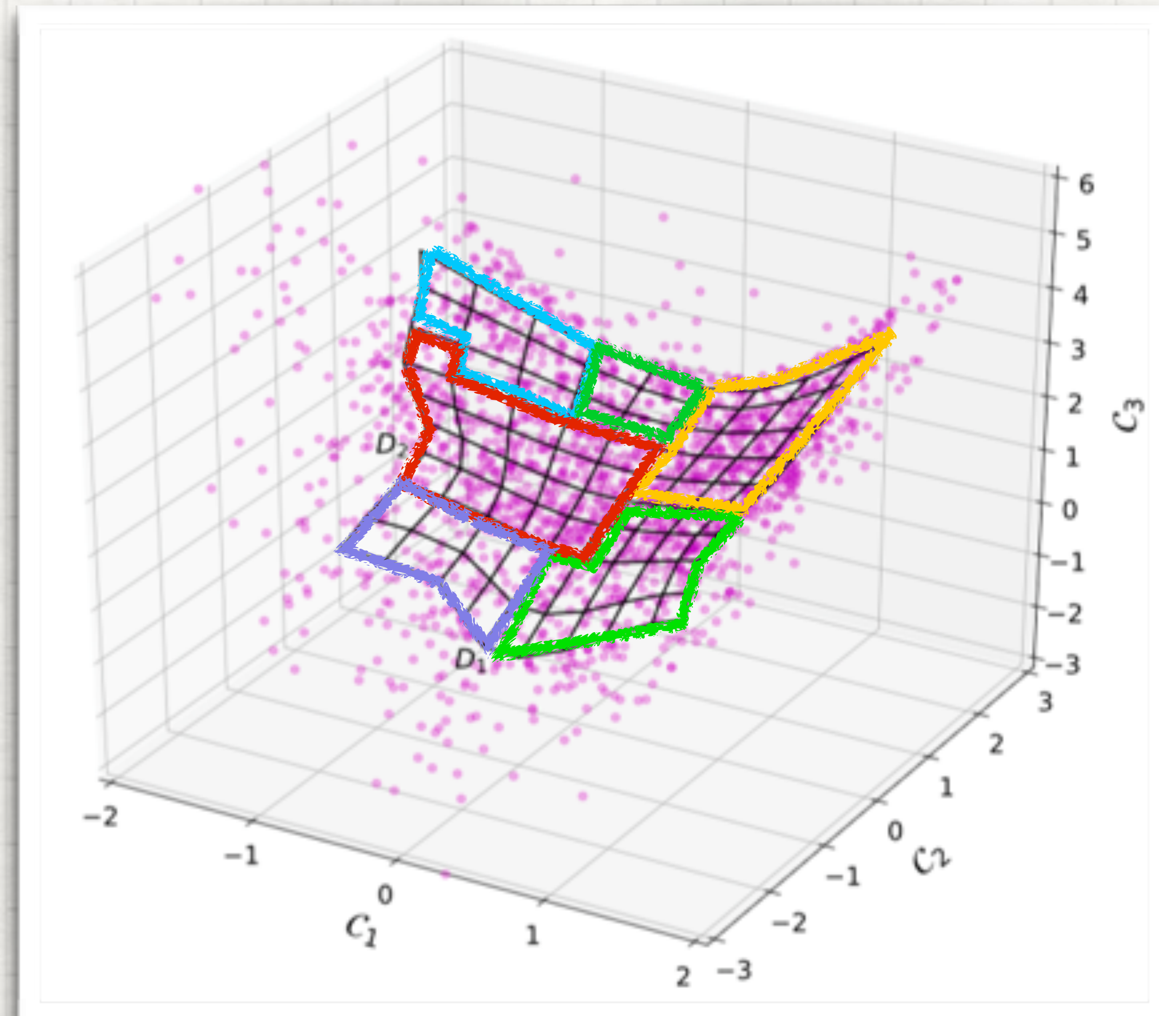


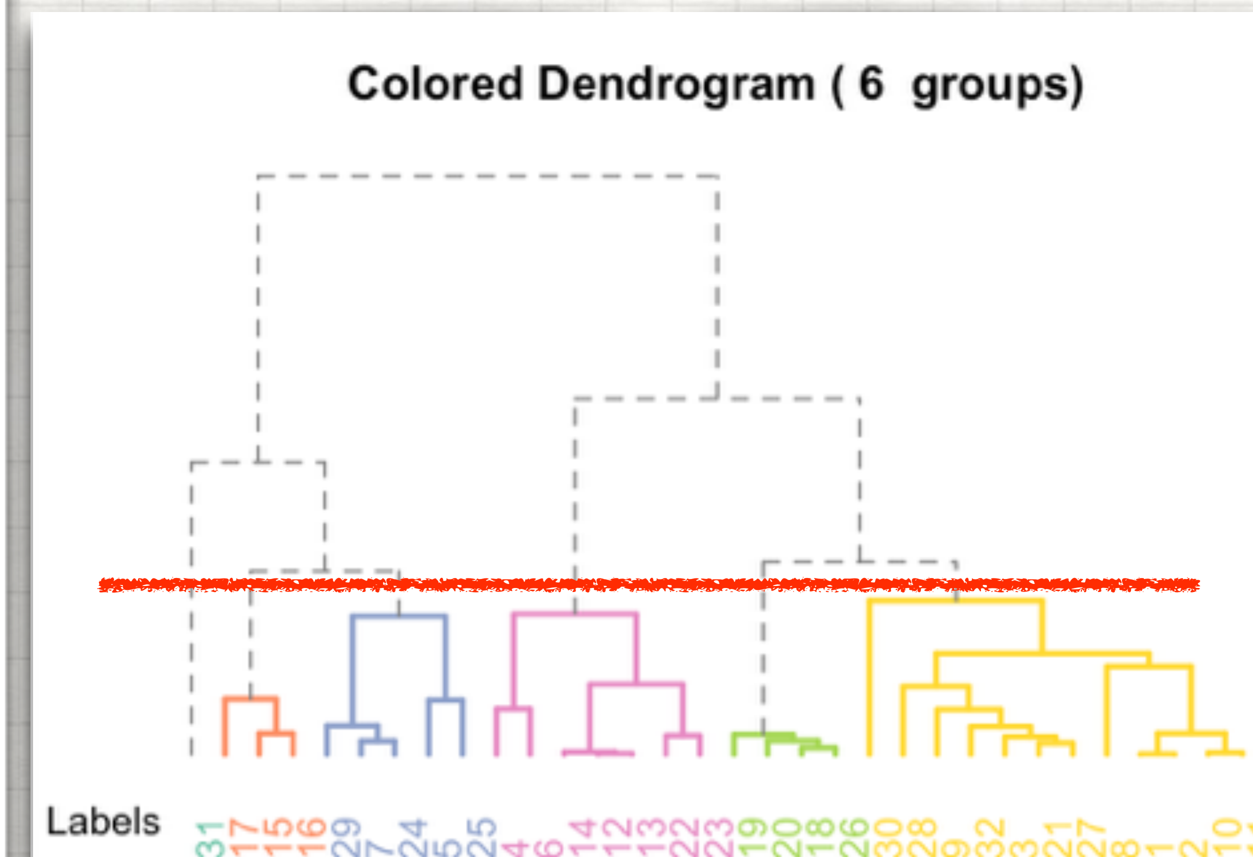
The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



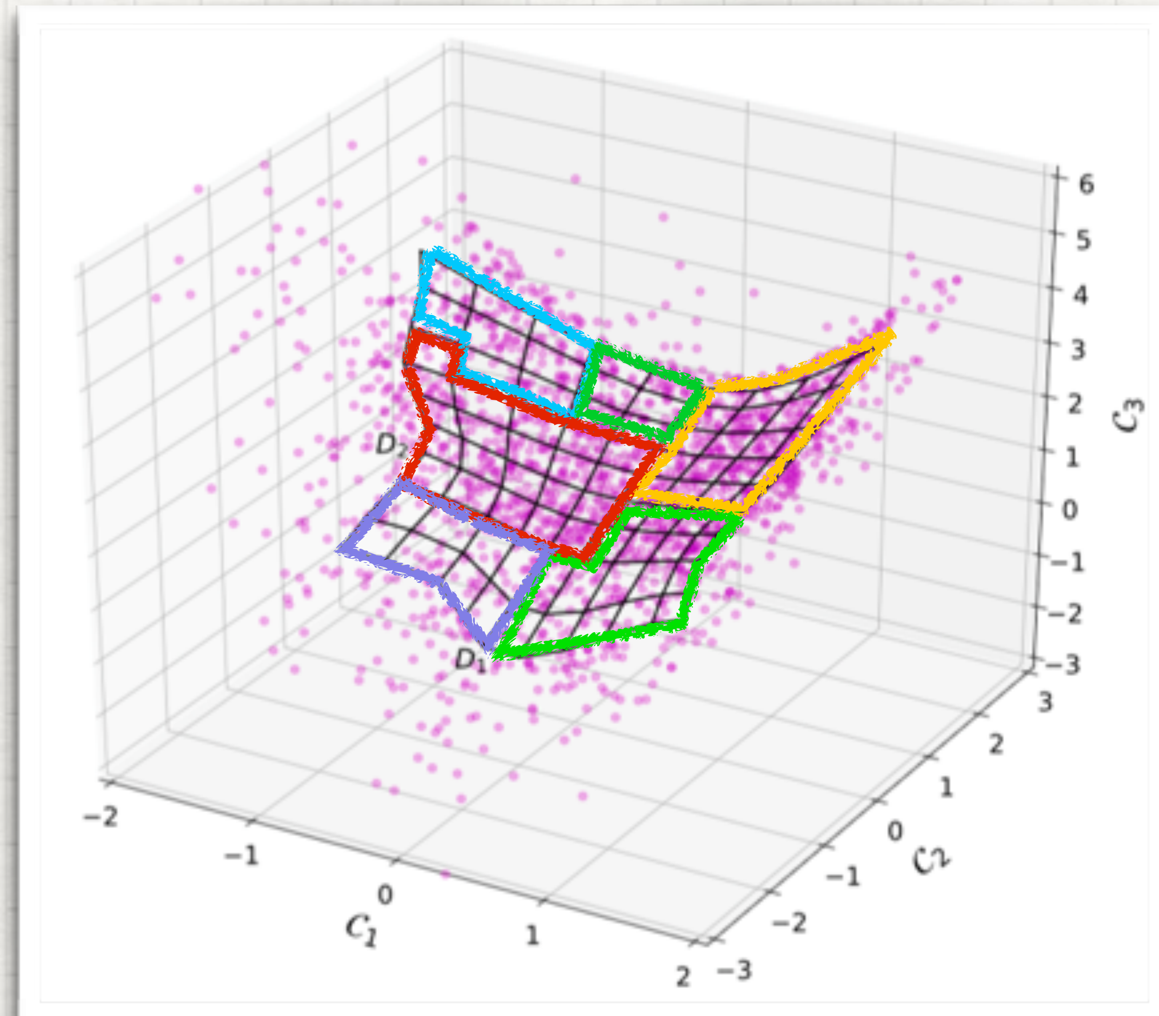


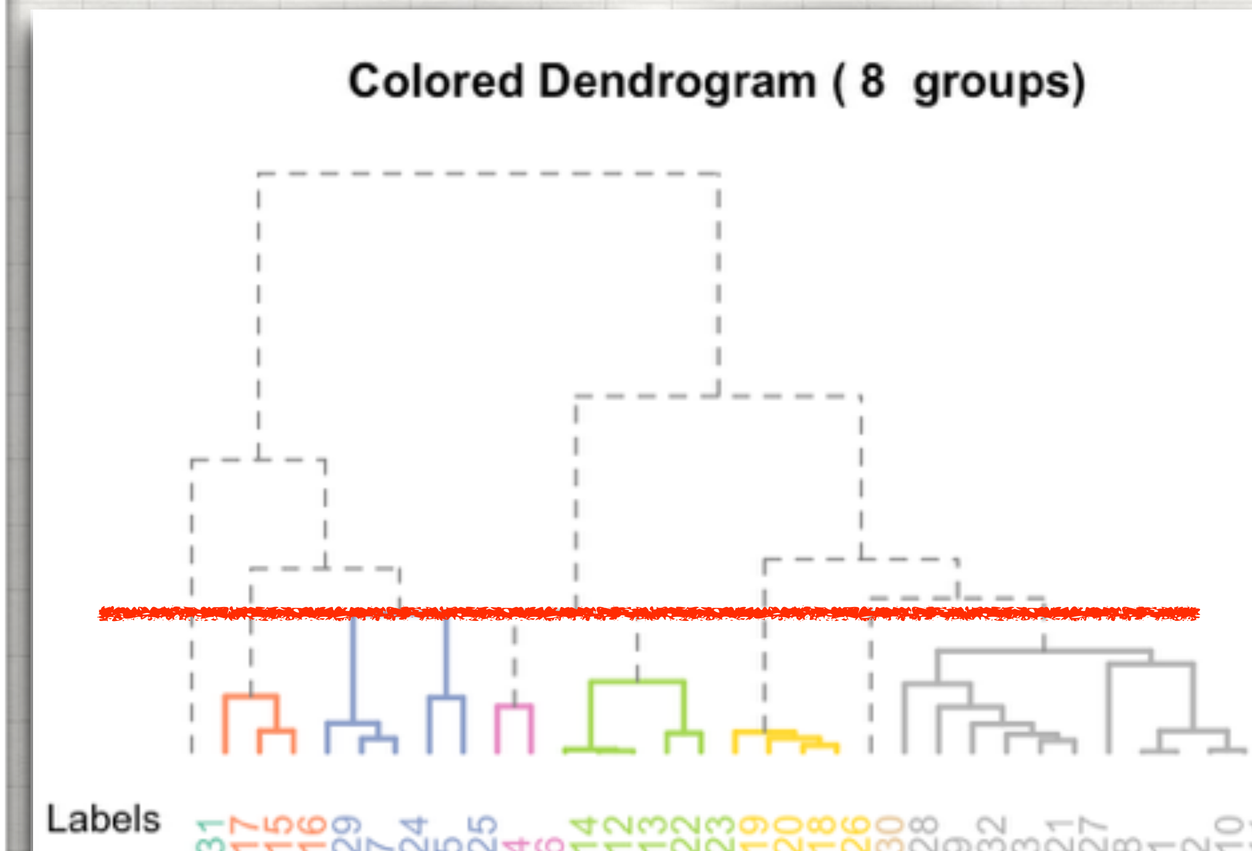
The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



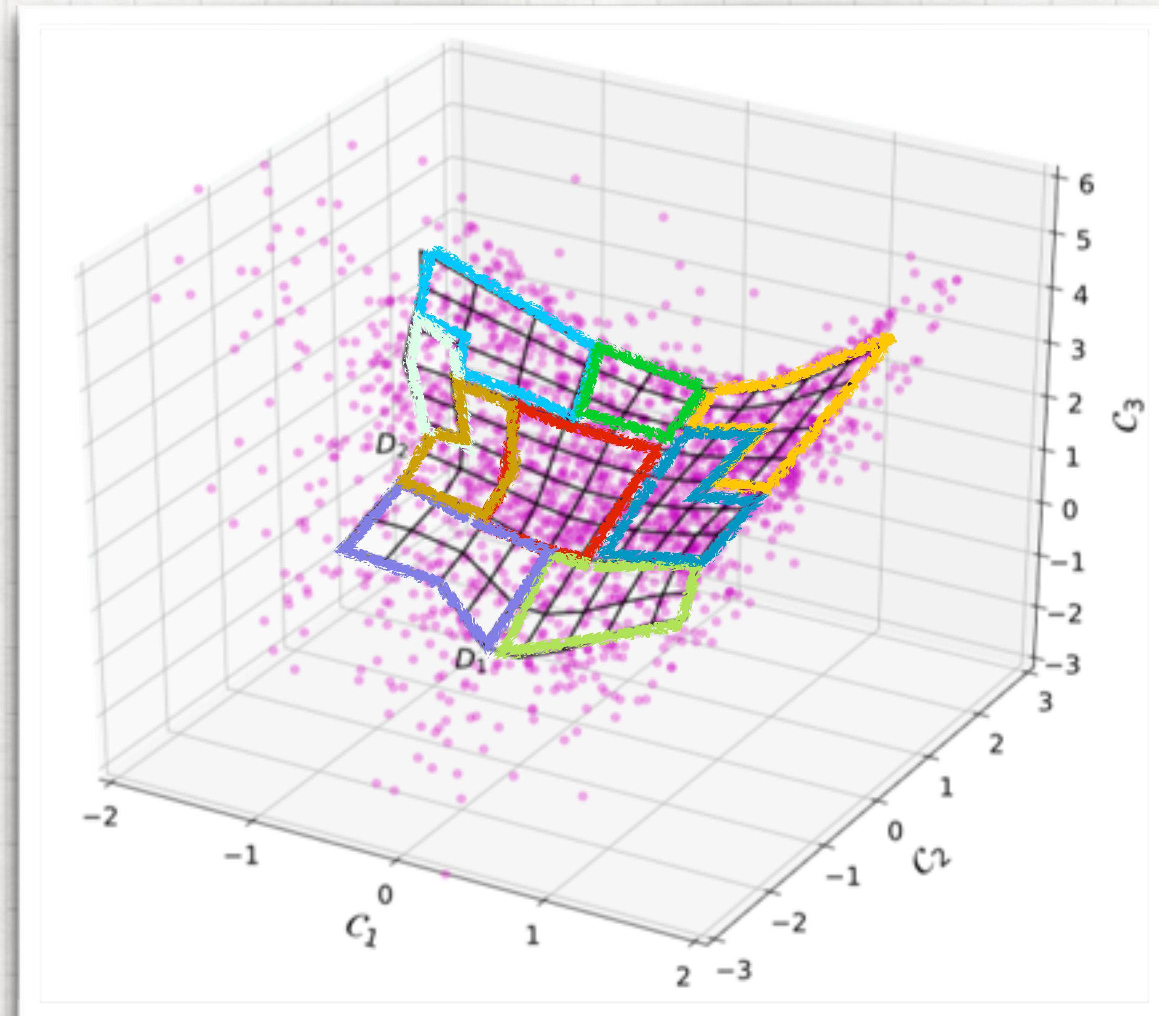


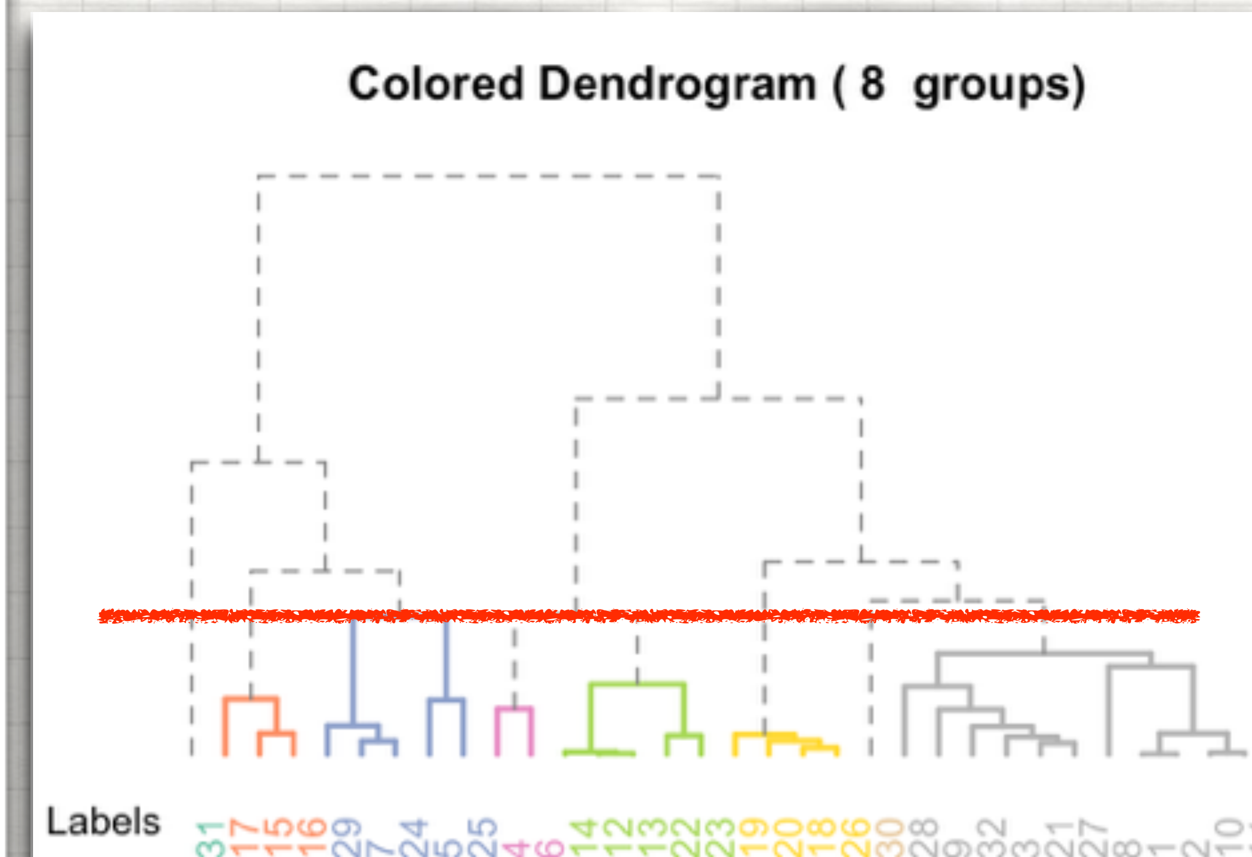
The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

CHOICE OF NEIGHBOUR DISTANCE/SOM PIXEL CLUSTERING



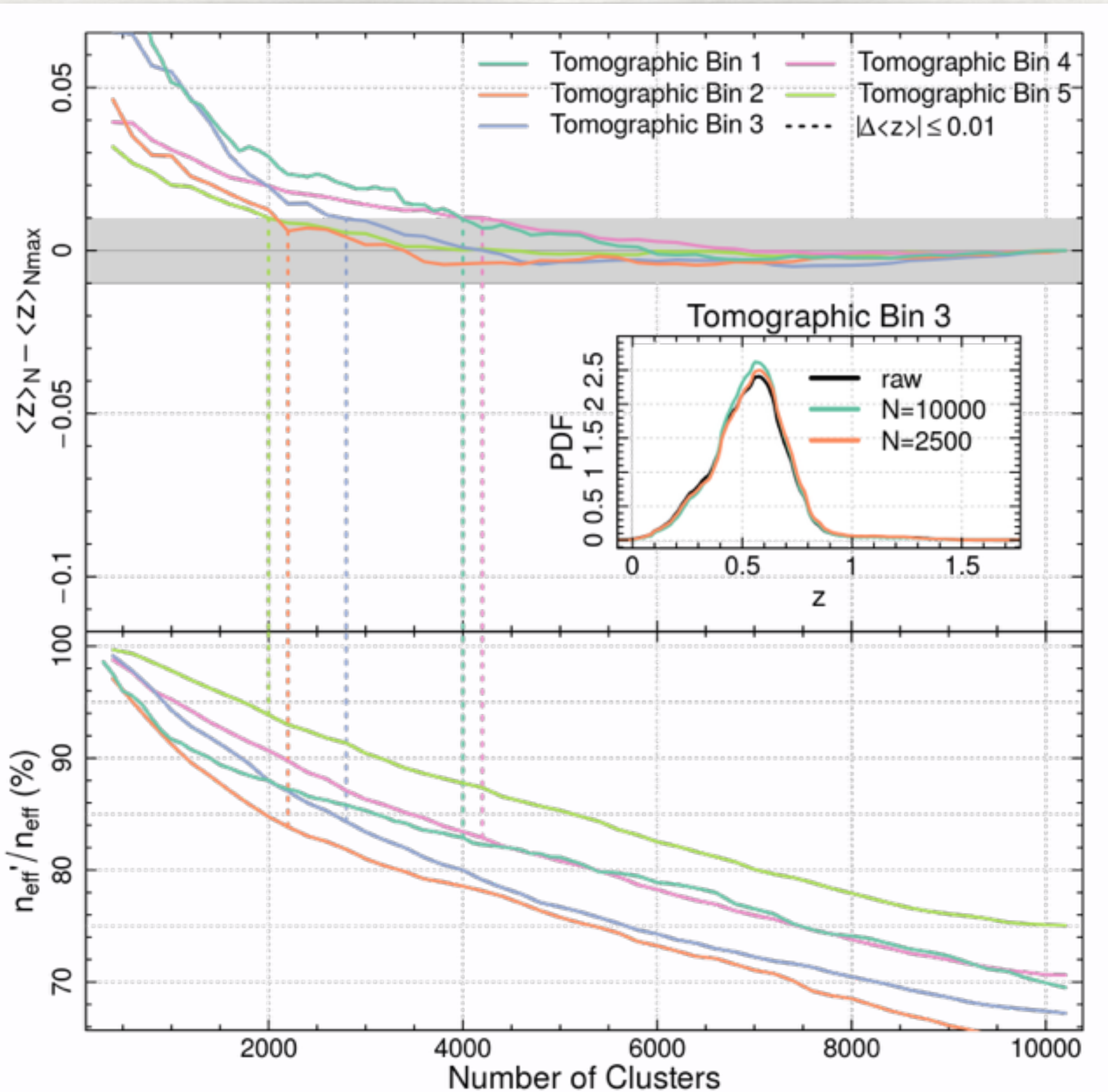


The choice of association number is then relevant.

↳An arbitrary number of associations can be constructed using hierarchical clustering of the SOM cells

Image credit: Davidson et al (2019)

SOM PIXEL CLUSTERING



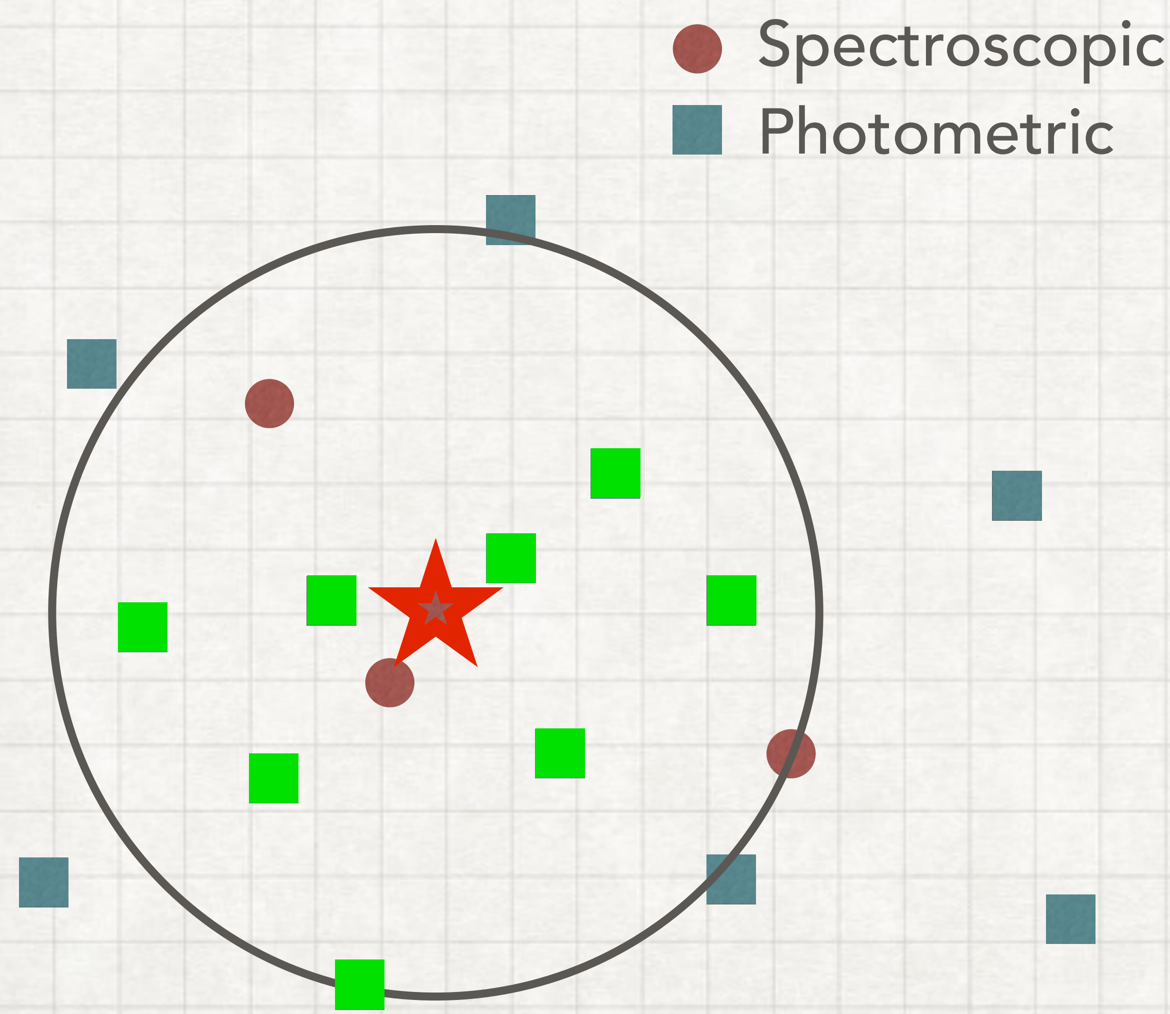
The goal is redshift calibration. So we can select the cluster number that produces unchanging mean-z results.

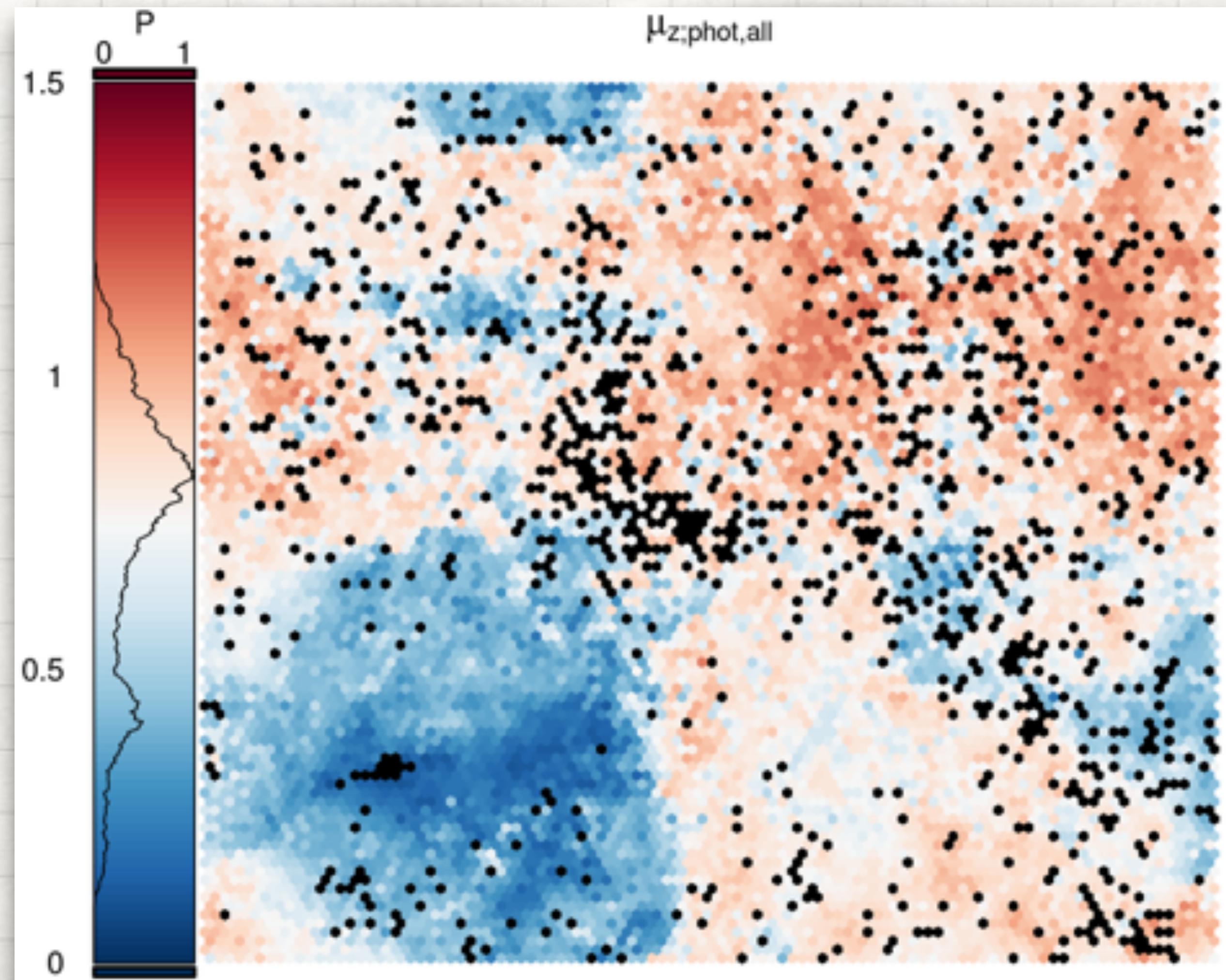
↳ mean-z is quite stable, neff is not

Wright et al (submitted)

REDSHIFT CALIBRATION REQUIREMENTS

The reweighting method requires that the associations have ~unique redshifts

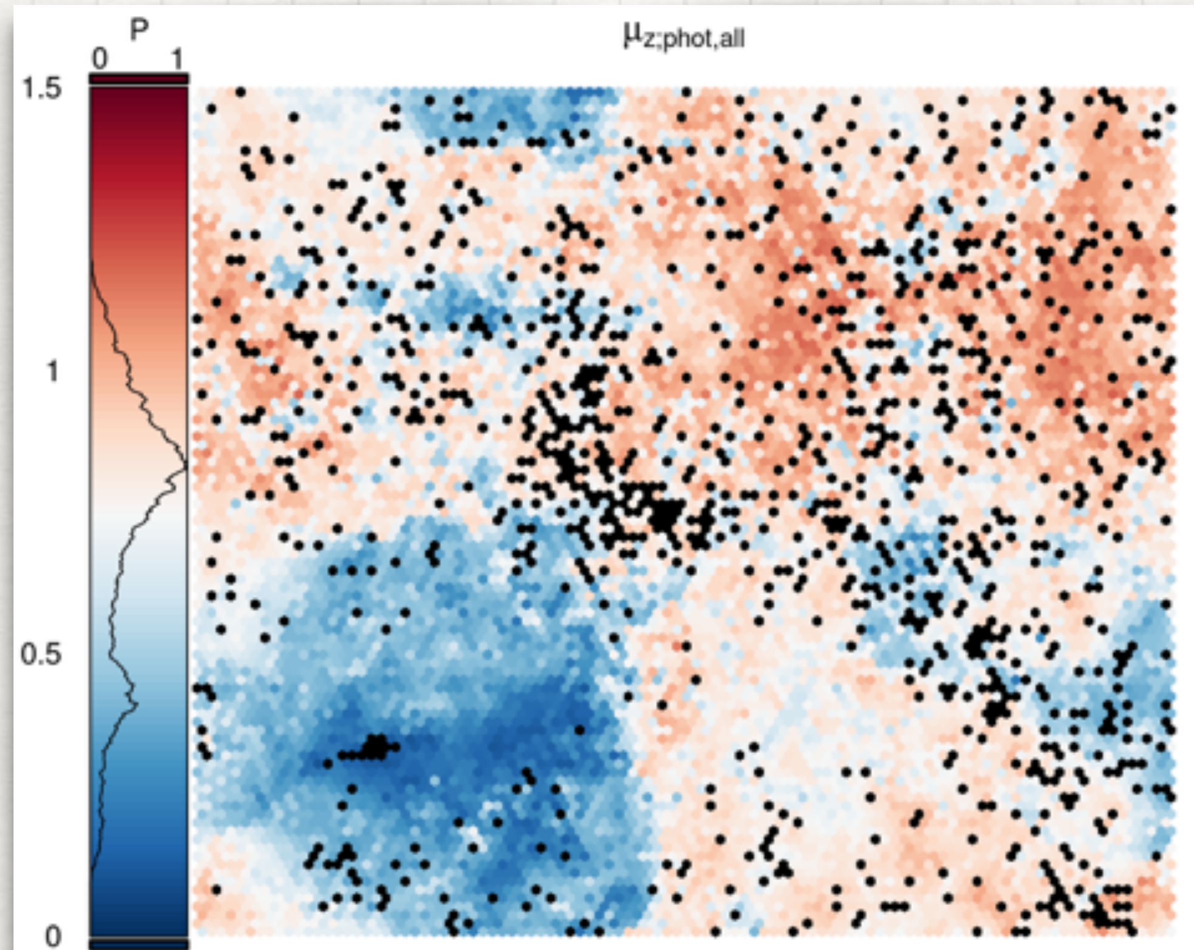


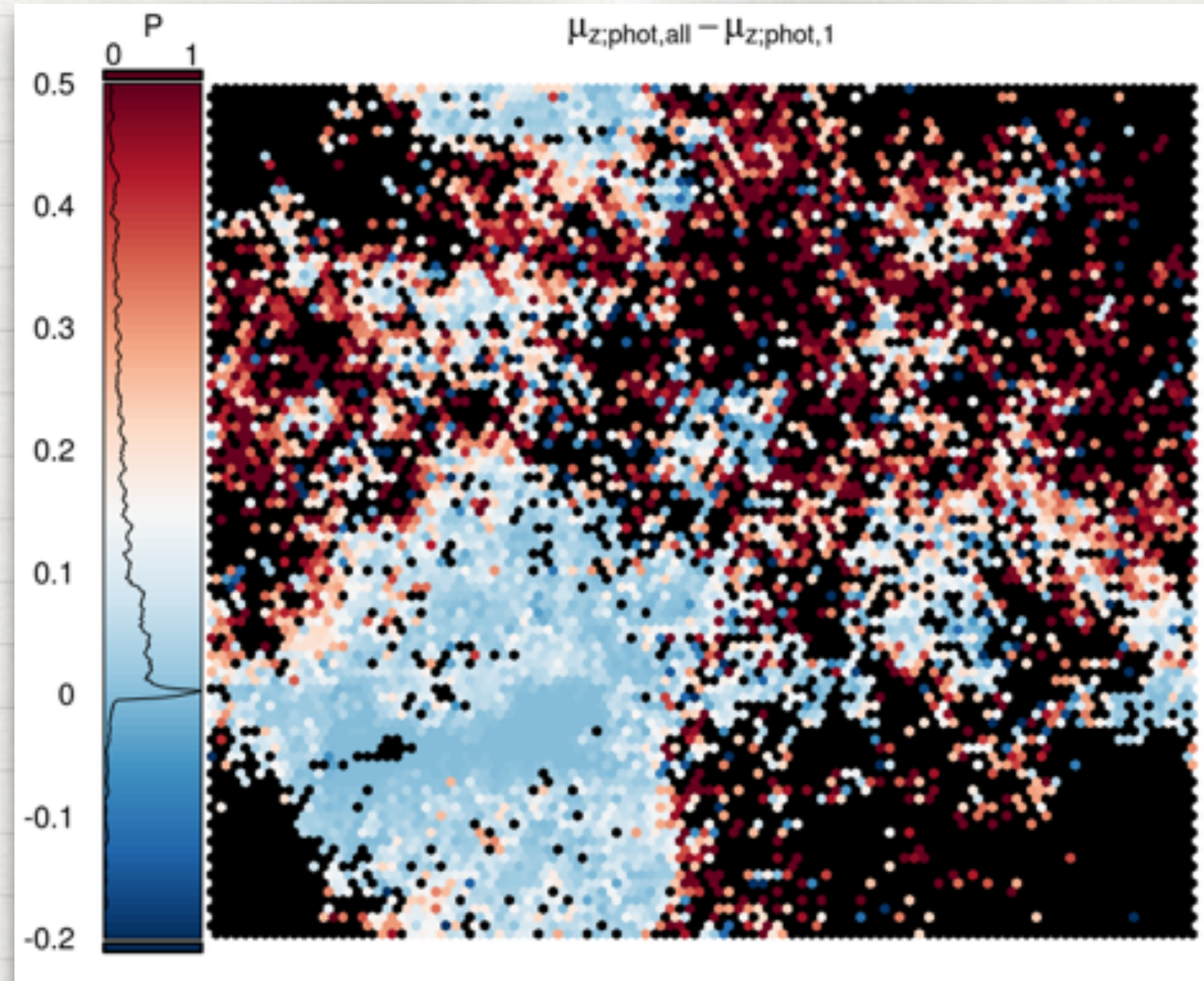


Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TOMOGRAPHIC BINNING



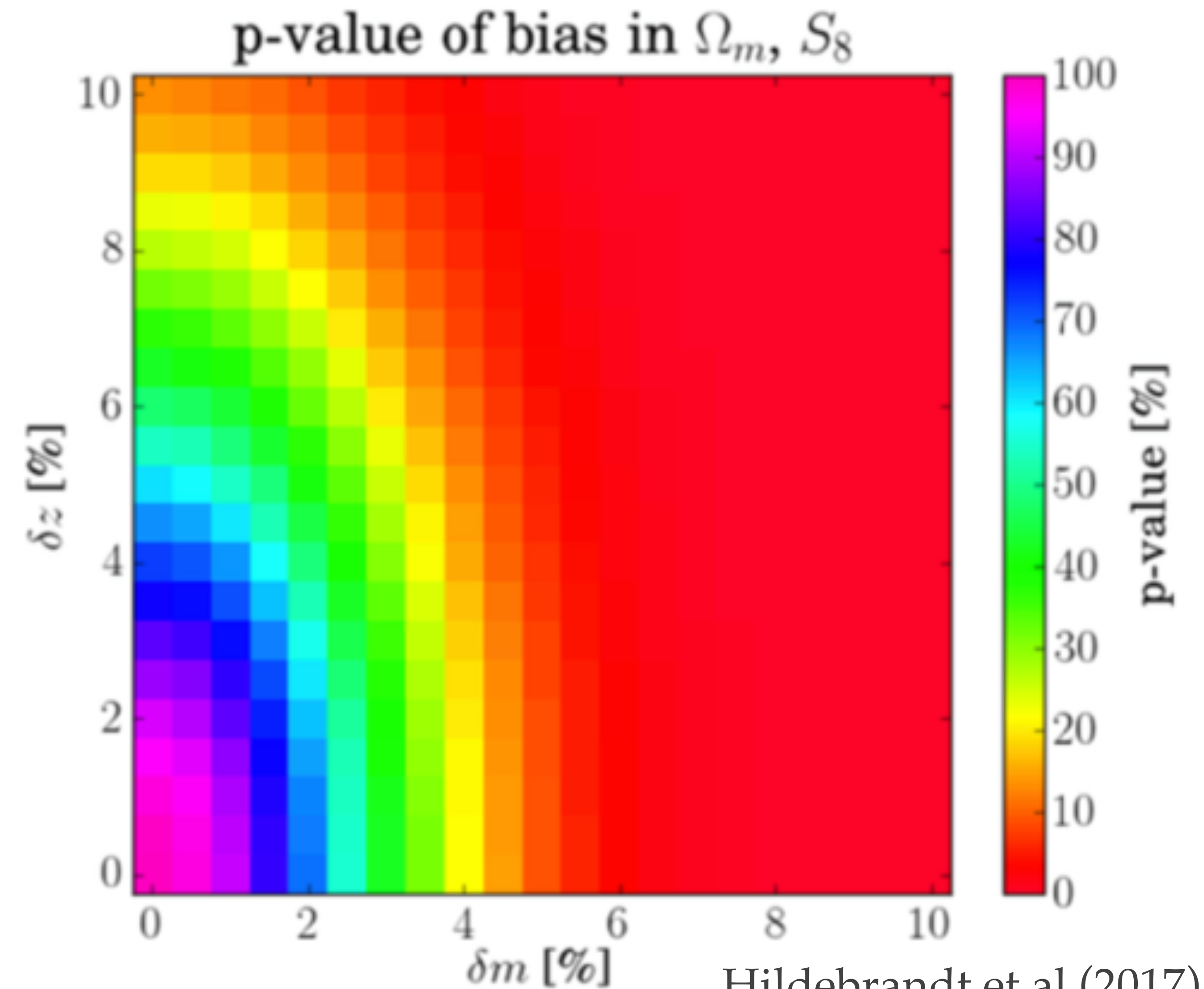


With tomographic binning, we apply a strong redshift-dependant selection effect. This modifies the N(z) of the SOM pixels, and necessitates like-for-like binning of the spectra.

Not previously implemented in kNN methods

Wright et al (submitted)

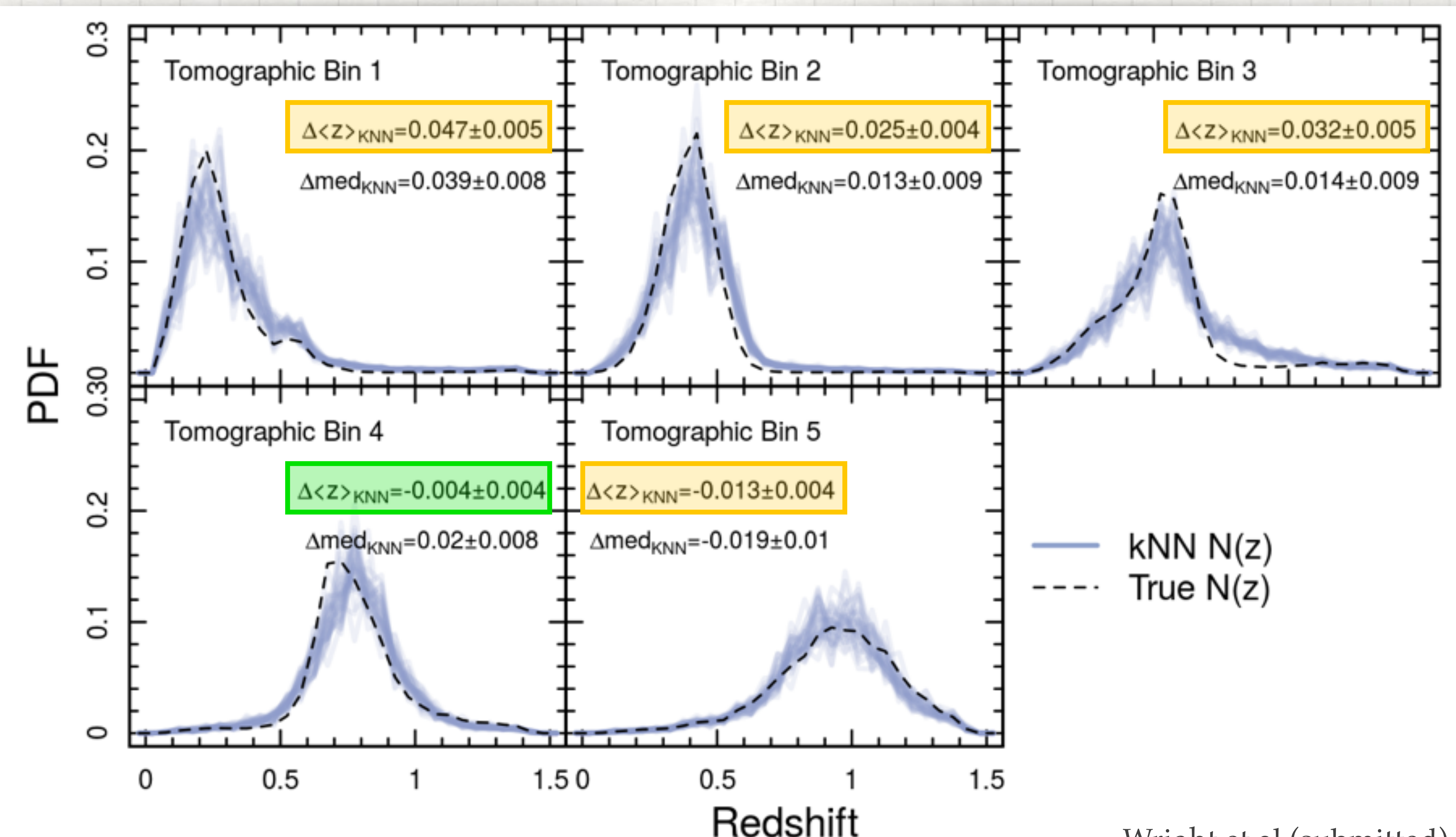
TESTING THE DIRECT CALIBRATION



Hildebrandt et al (2017)

TESTING THE DIRECT CALIBRATION

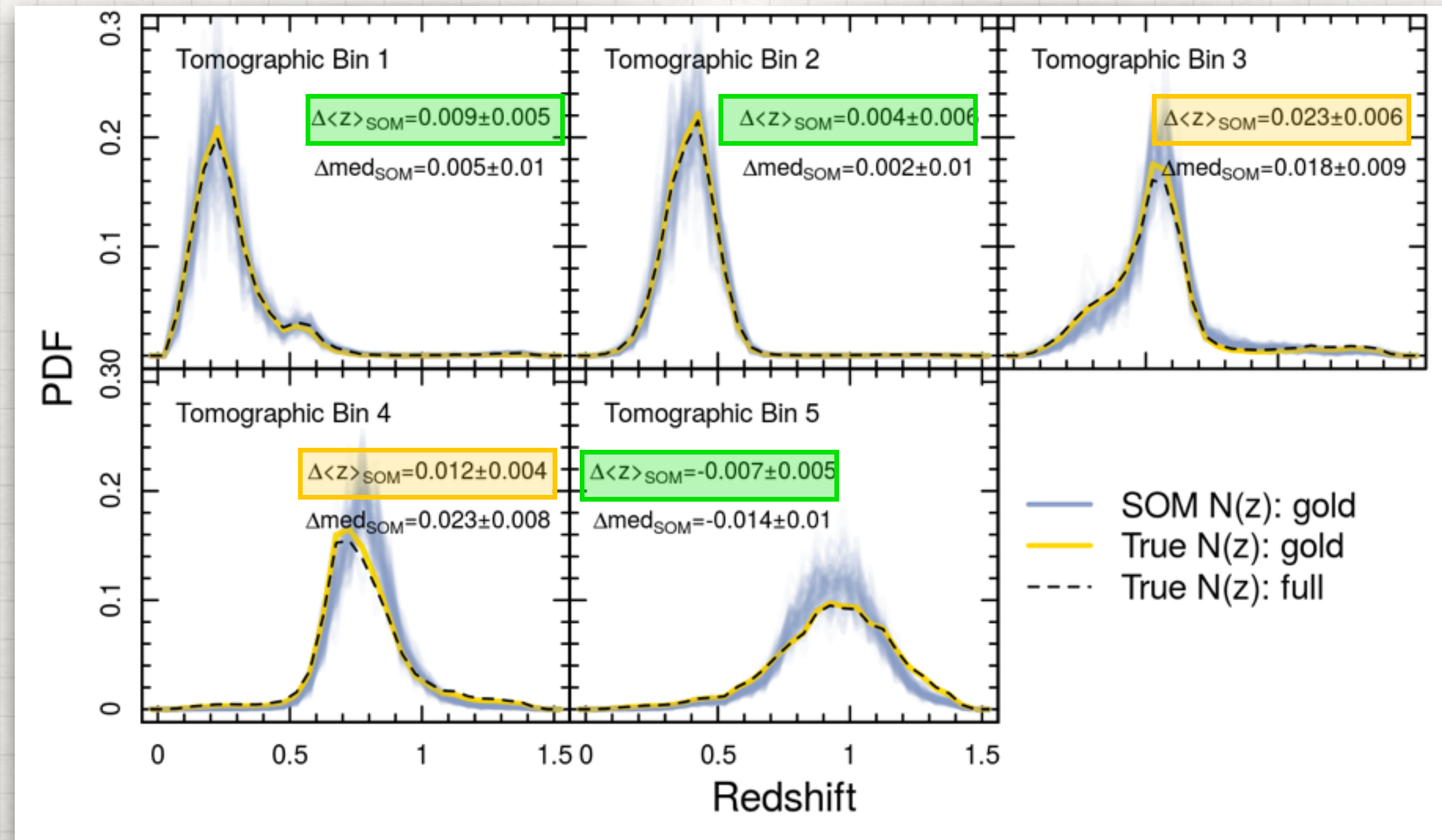
100 spectroscopic lines-of-sight & KiDS-like photometry using MICE2



Wright et al (submitted)

SOM PHOTOMETRIC REDSHIFT CALIBRATION

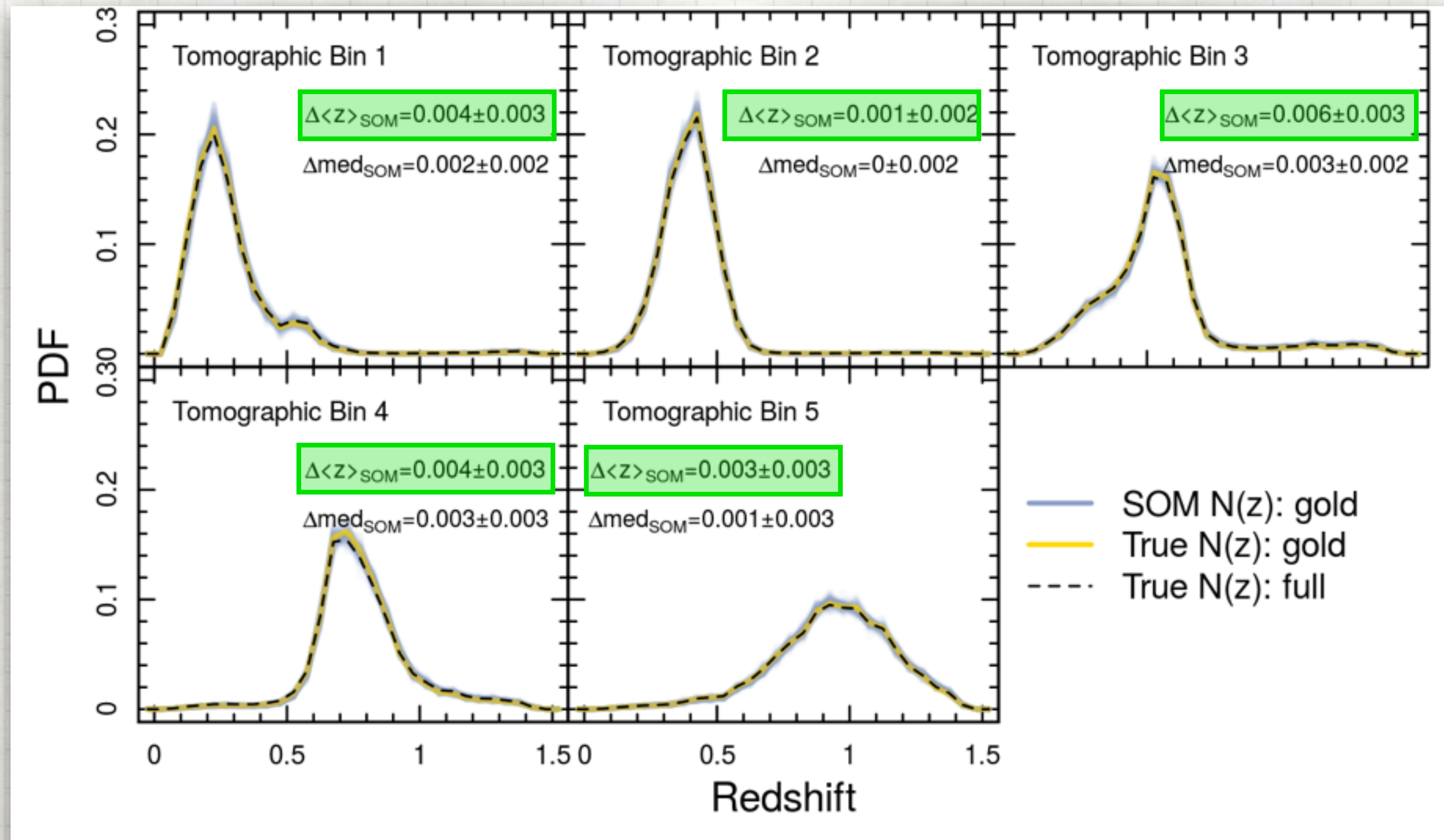
KIDS-LIKE DATA



Wright et al (submitted)

SOM PHOTOMETRIC REDSHIFT CALIBRATION

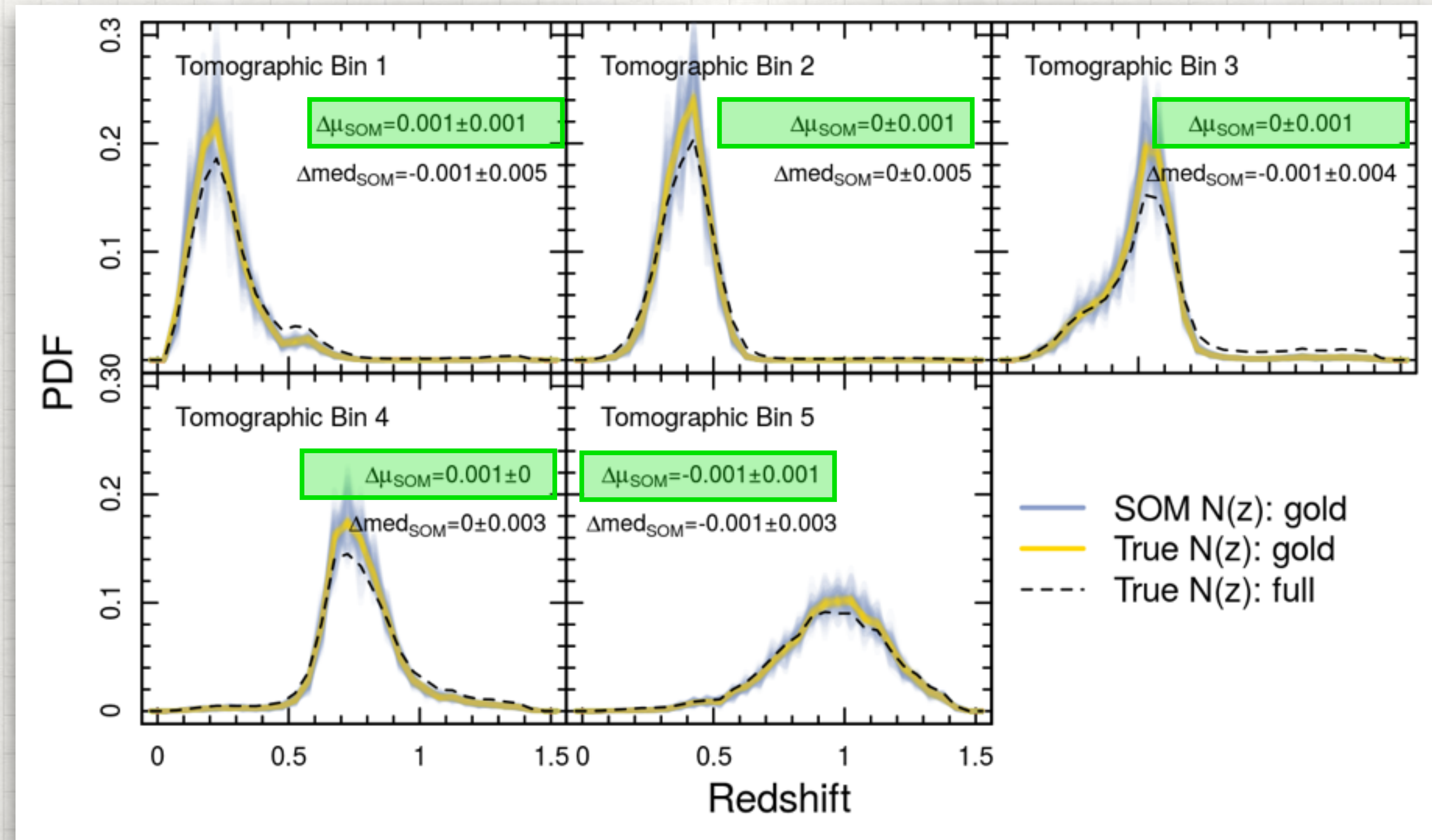
PERFECT SPECTROSCOPY, NOISY PHOTOMETRY



Wright et al (submitted)

SOM PHOTOMETRIC REDSHIFT CALIBRATION

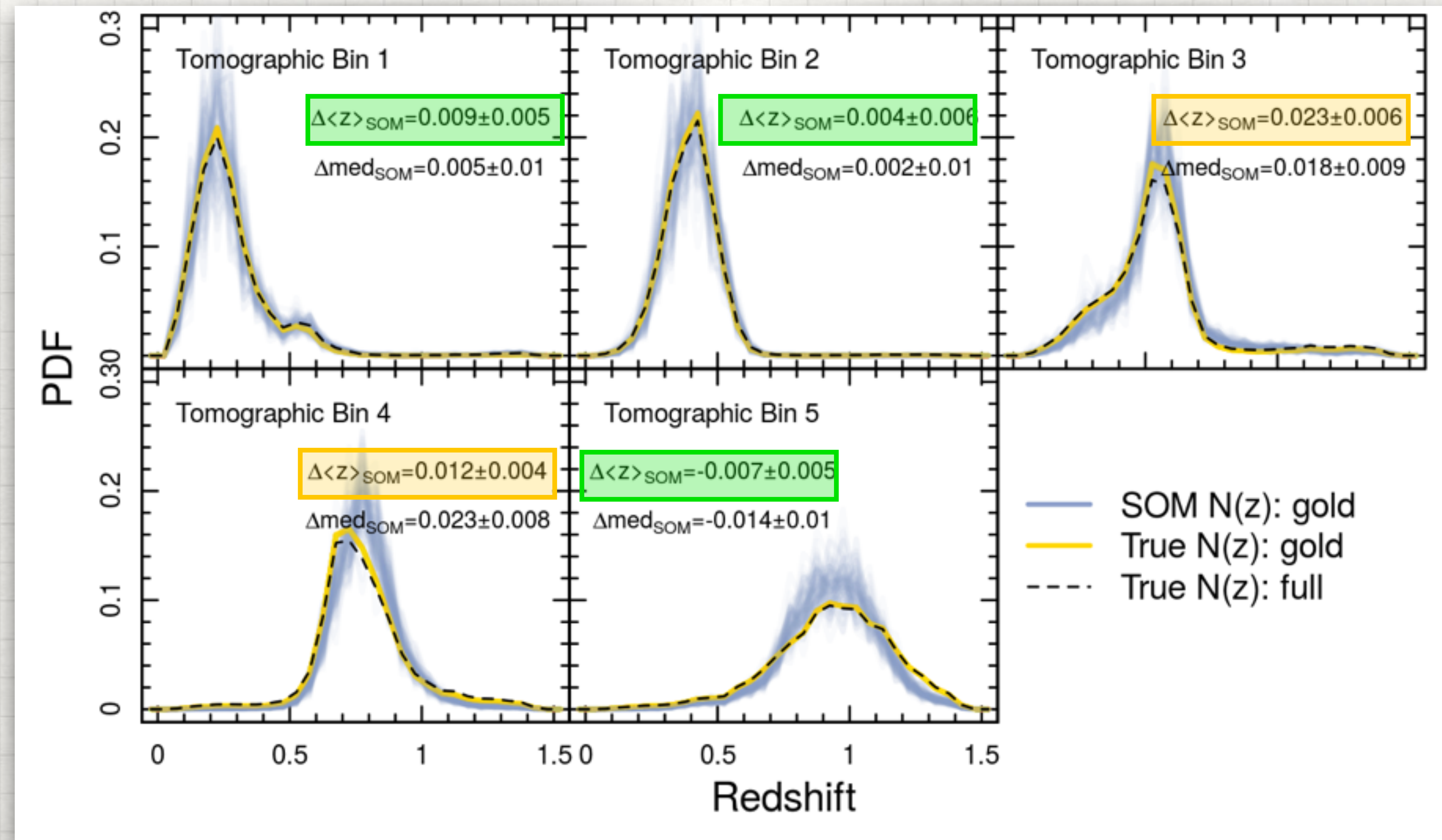
BIASED SPECTROSCOPY, NOISELESS PHOTOMETRY



Wright et al (submitted)

SOM PHOTOMETRIC REDSHIFT CALIBRATION

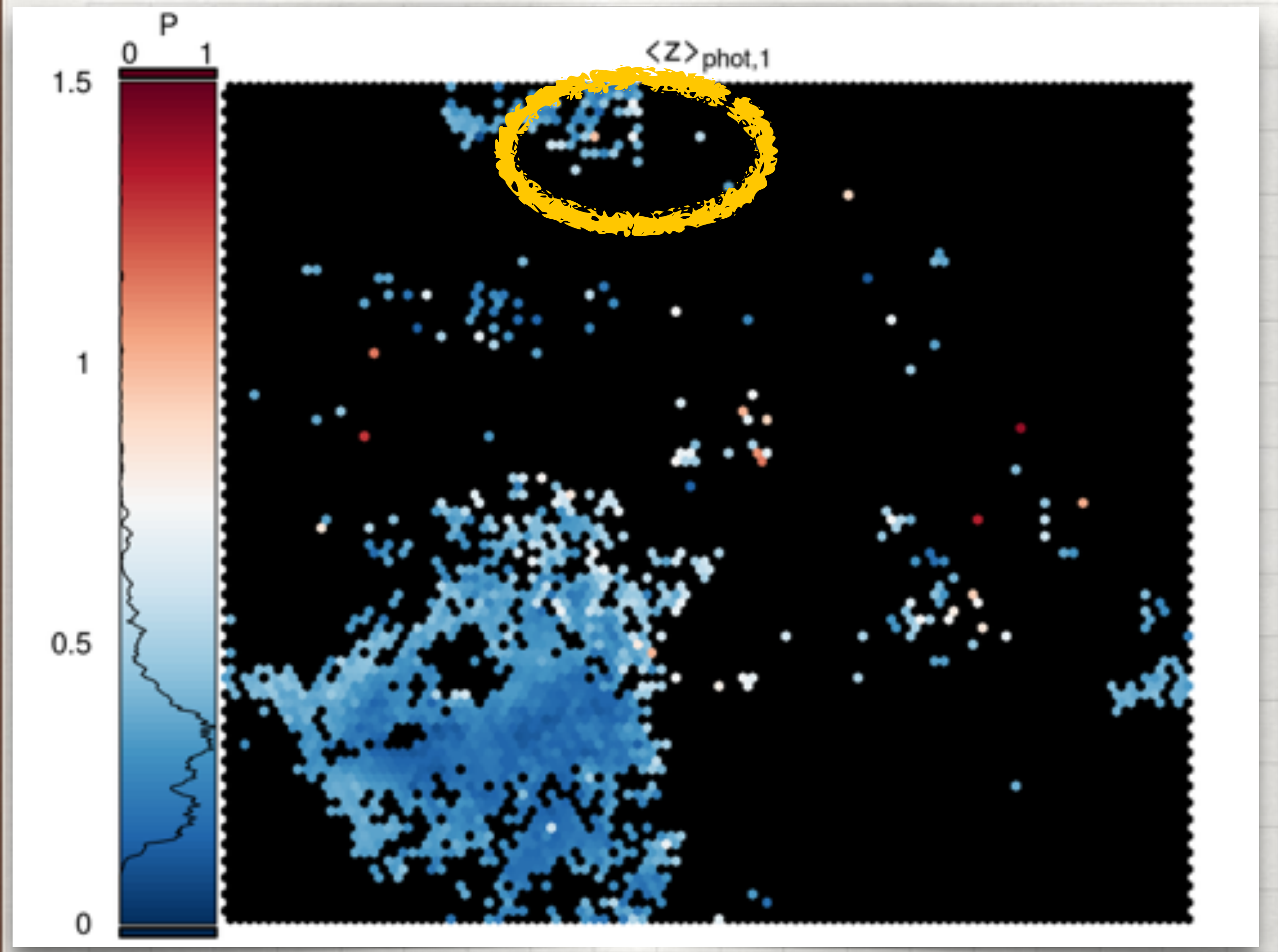
KIDS-LIKE DATA

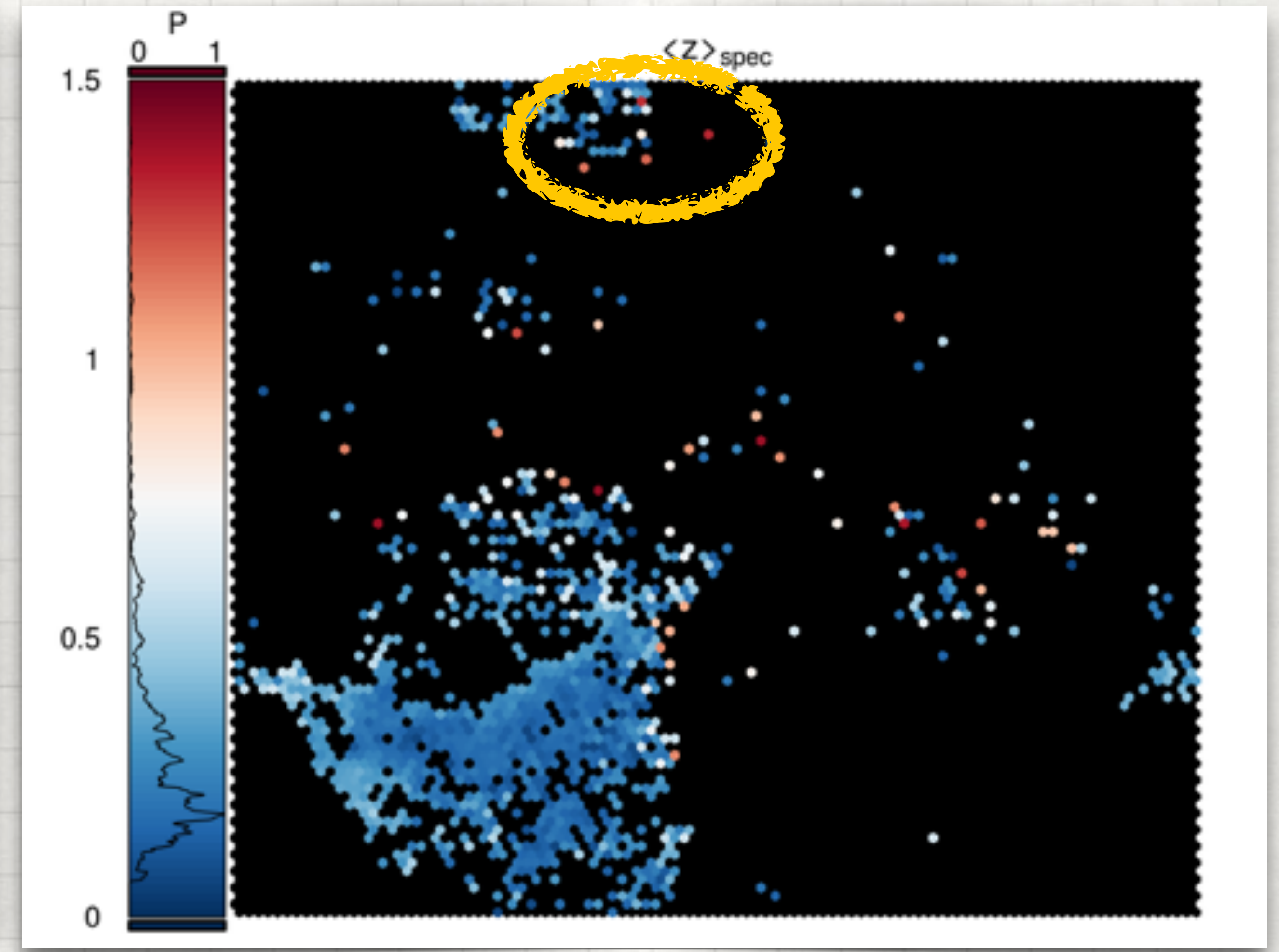


Wright et al (submitted)

IMPROVING THE CALIBRATION FURTHER

QUALITY CONTROL





The bias in the noisy+spectroscopically selected calibration comes from a small number of biased associations. These can be flagged and removed.

Wright et al (submitted)

FINAL CALIBRATION STATISTICS

Dataset Type	Phot	Reconstruction Bias ($\Delta\langle z\rangle$) bin1	bin2	bin3	bin4	bin5
perfect	exact	$0.001\pm <10^{-3}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
KV450	exact	0.002±0.001	0.002±0.002	0.003±0.001	0.003±0.001	–0.001±0.001
Perfect	noisy	0.004±0.003	$<10^{-3}\pm0.002$	0.006±0.003	0.004±0.003	0.003±0.003
KV450	noisy	0.009±0.005	0.004±0.006	**0.023±0.006**	**0.012±0.004**	–0.007±0.005
KV450	noisy+QC1	$<10^{-3}\pm0.005$	0.002±0.006	**0.013±0.006**	**0.011±0.004**	–0.006±0.005
KV450	noisy+QC2	0.002±0.005	0.003±0.006	0.007±0.005	0.009±0.004	–0.006±0.004
				using kNN association		
KV450	noisy	**0.047±0.005**	**0.025±0.004**	**0.032±0.005**	–0.004±0.004	**–0.013±0.004**

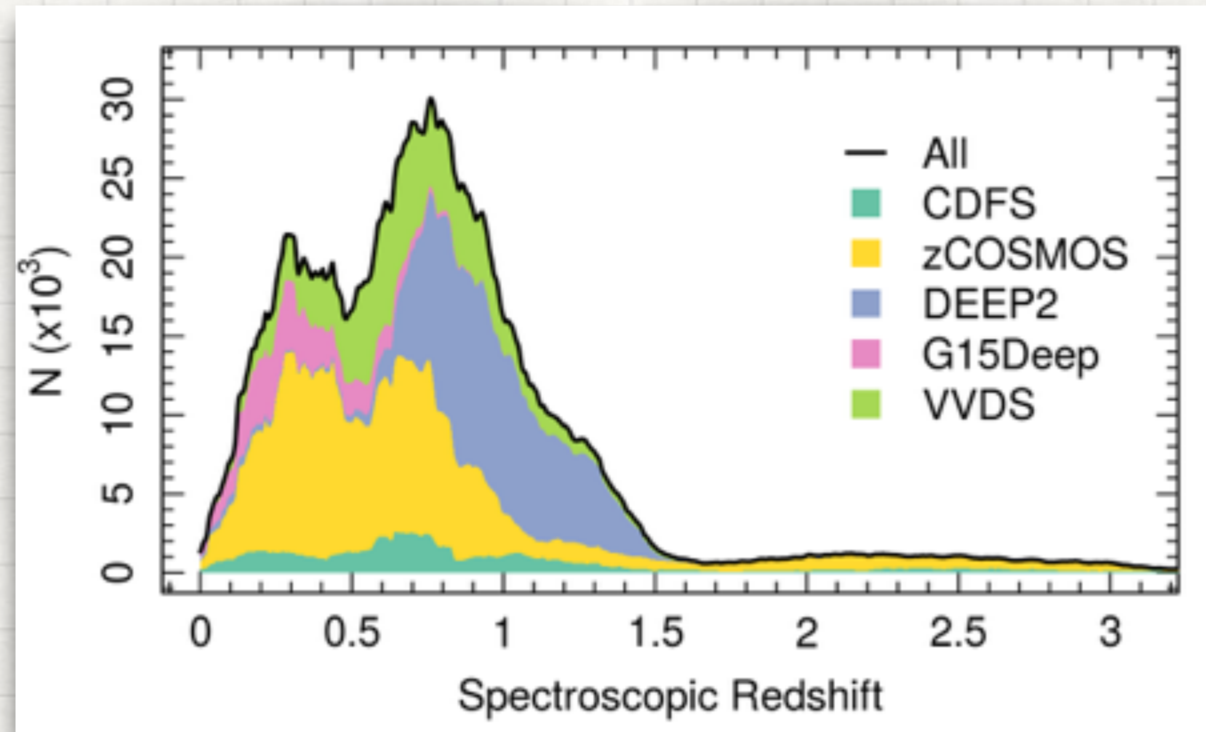
The updated calibration produces more accurate calibration, and is able to reach ~0.01 ± 0.01 precision in all bins when performing some quality control.

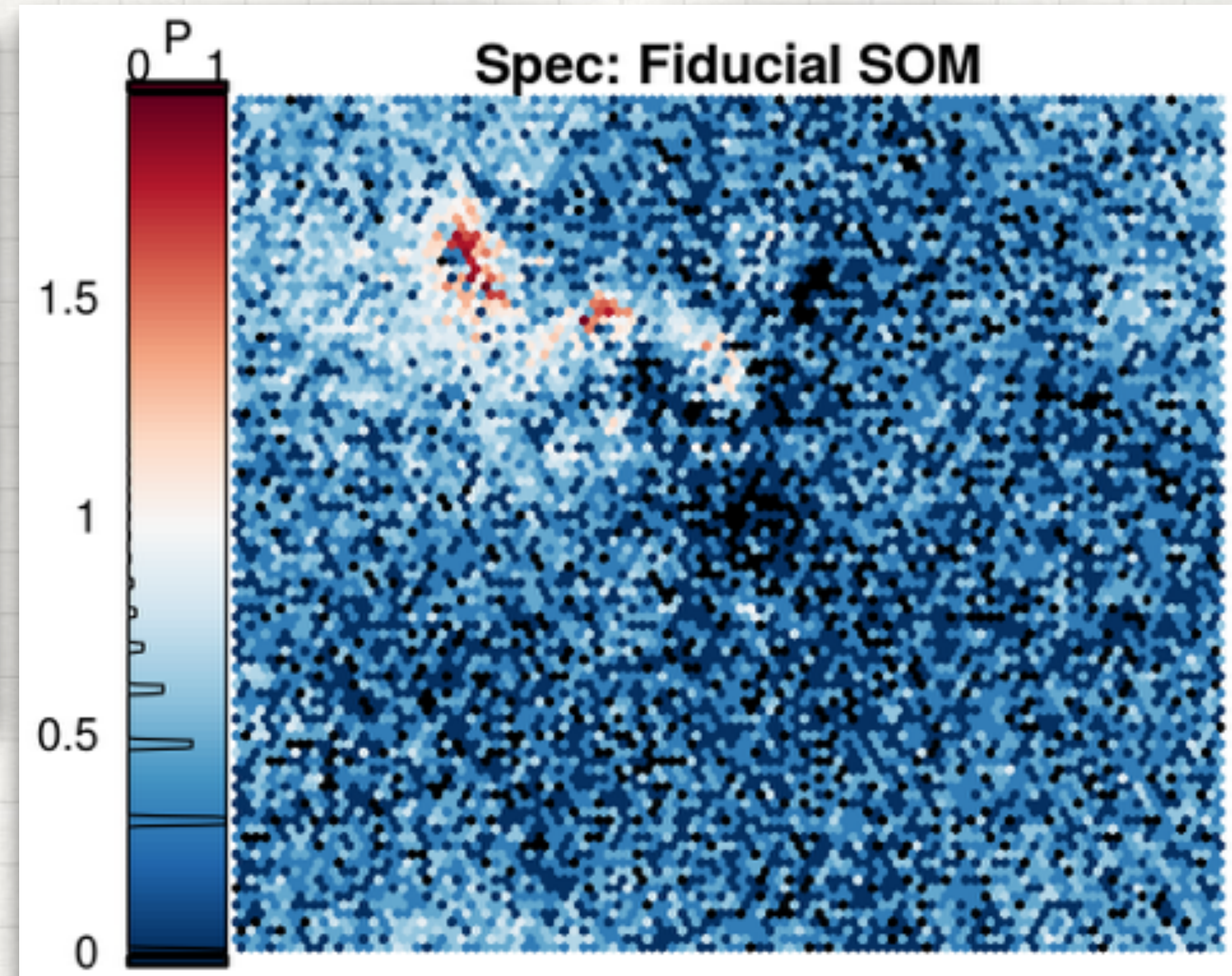
The interpretation is that this reduction in bias comes from the removal of non-represented photometric data

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TESTING SPECTRAL DATASETS





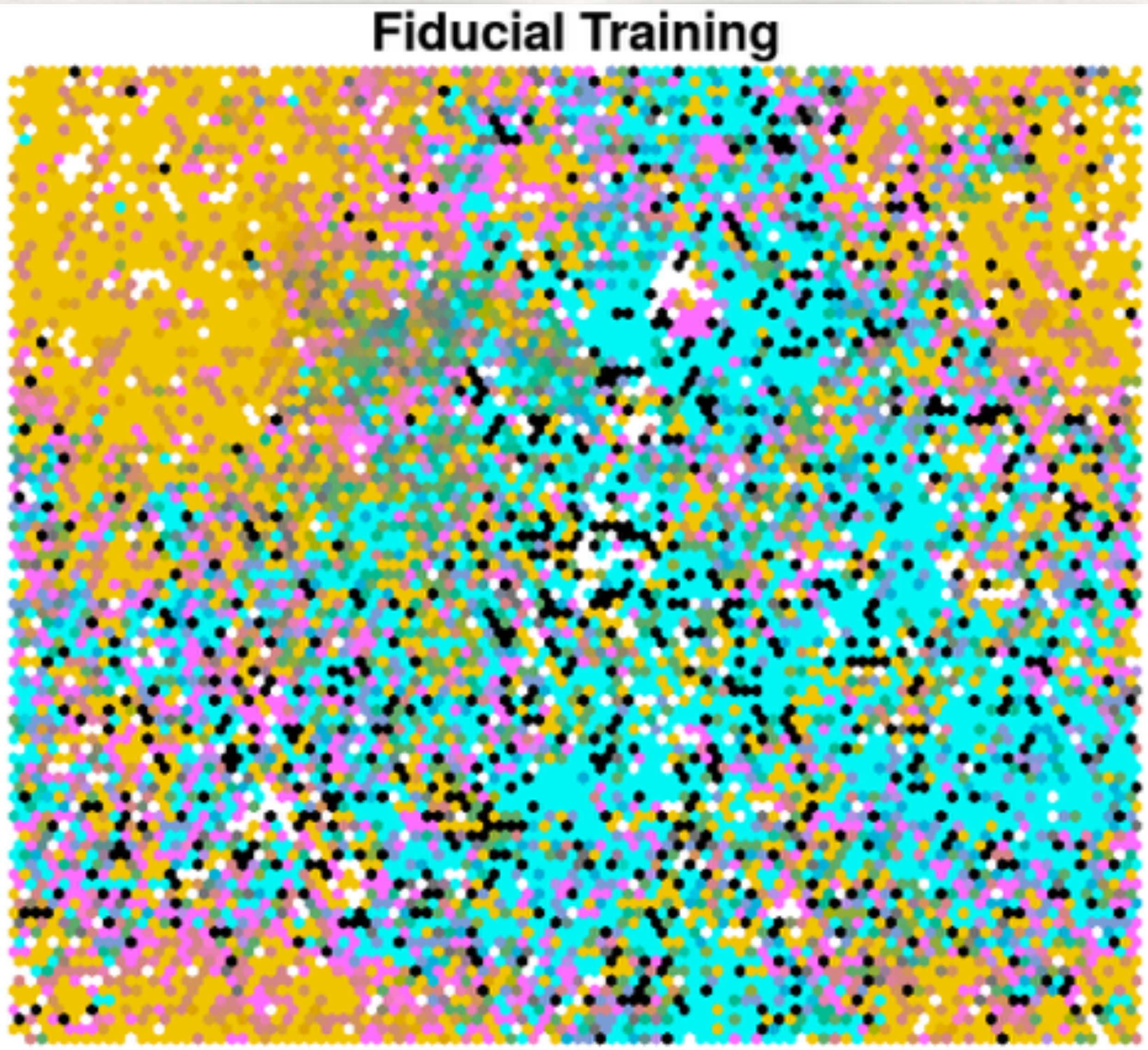
The cosmic shear spectroscopic compilation is an amalgam of many different spectroscopic datasets

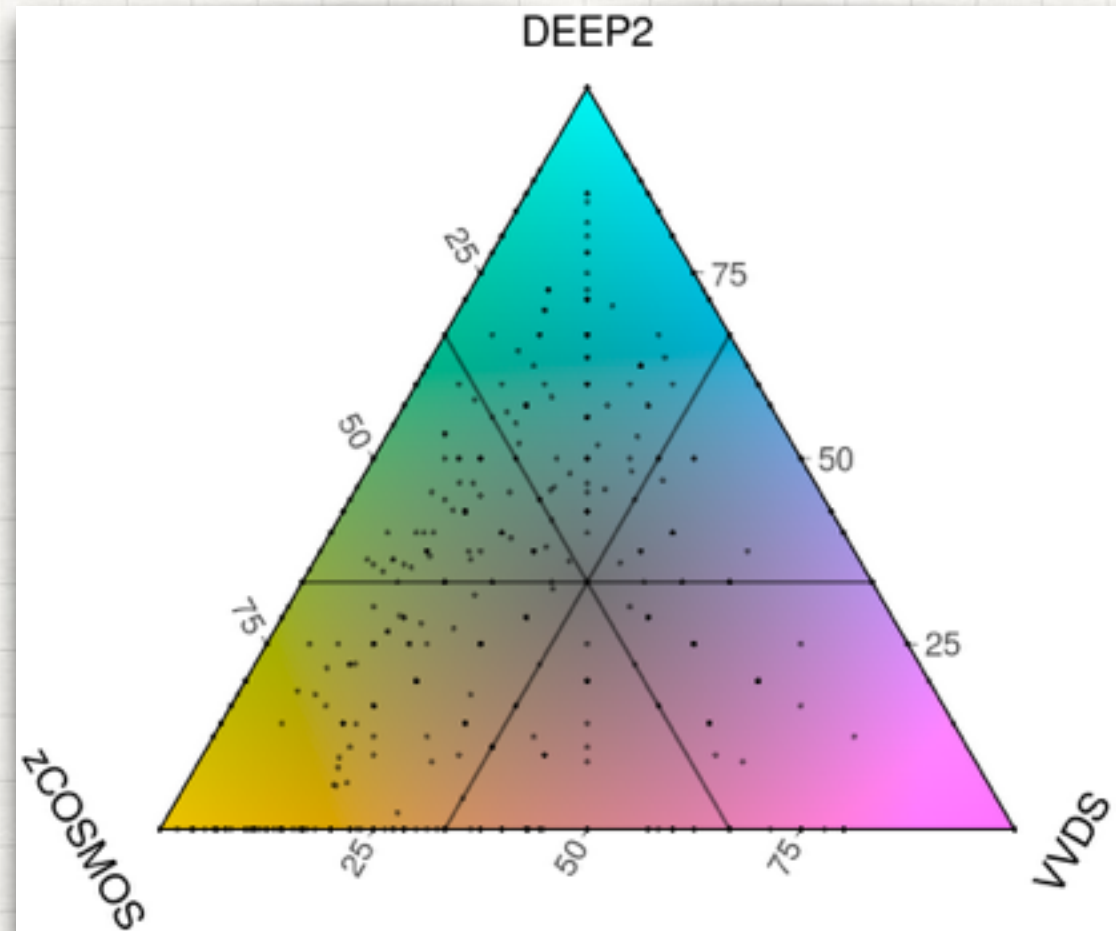
This dataset has a complicated colour mapping, and various selection biases. How representative is our compilation? Does any one survey dominate our calibration?

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TESTING SPECTRAL DATASETS



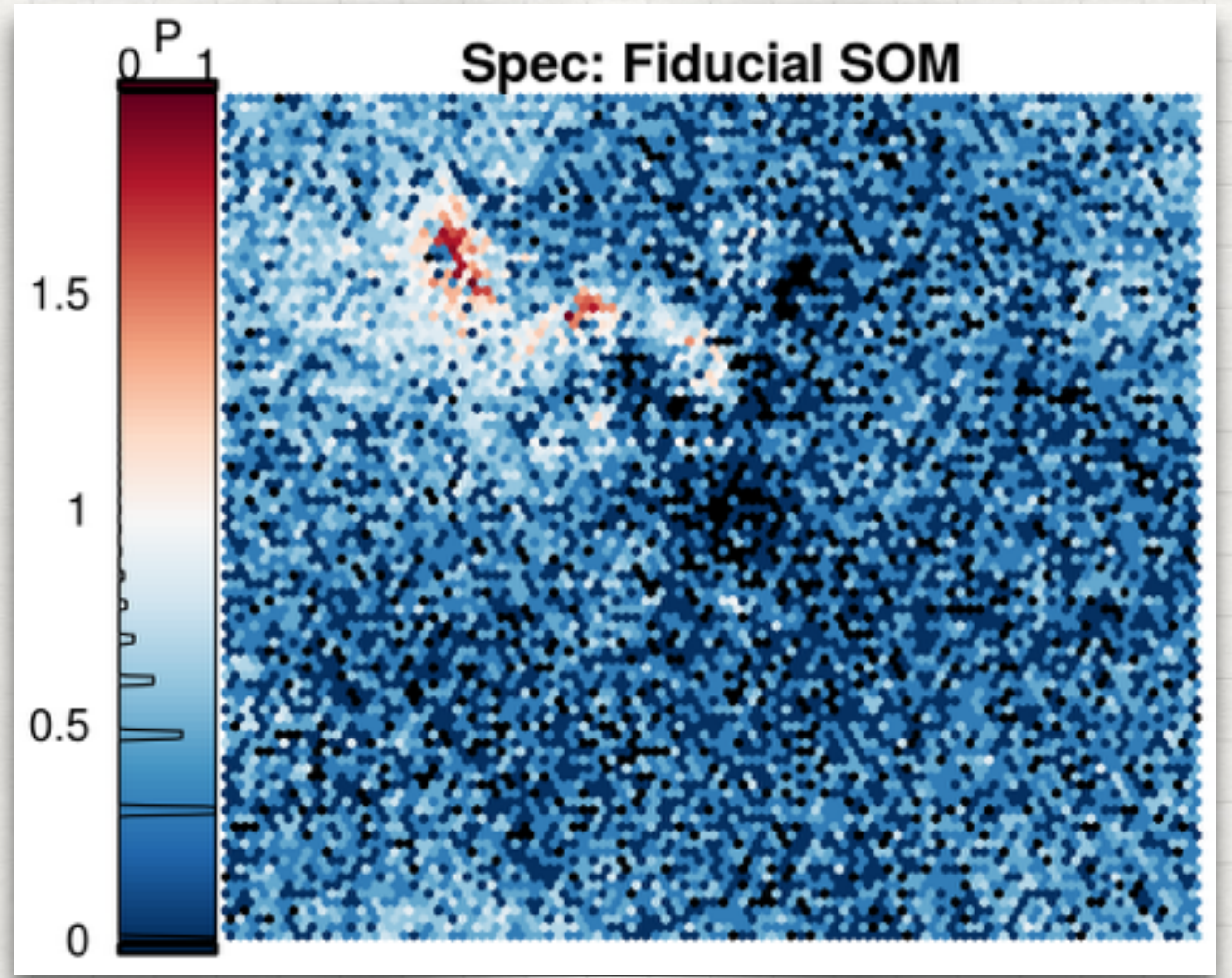


The spectroscopic data are highly complementary: rarely do two or more datasets occupy any one SOM pixel

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TESTING SPECTRAL DATASETS



9% of our spectroscopic maps pixels are devoid of spectra.

However these pixels host only 0.5% of the 2D cosmic shear photometric data.

↳ Pixels without spectroscopy are overwhelmingly in insignificant parts of the colour-space for KiDS.

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TOMOGRAPHIC BINNING

Spectroscopic Compilation	Training Size (all)	$f_{\rm pix}$ (all, %)	$n'_{\rm eff}/n_{\rm eff}(\%)$ All $z_B \in [0.1, 1.2)$	bin1 $[0.1, 0.3)$	bin2 $[0.3, 0.5)$	bin3 $[0.5, 0.7)$	bin4 $[0.7, 0.9)$	bin5 $[0.9, 1.2)$
Full Sample	25373	91.9	99.5	82.7	83.9	84.6	82.7	94.0
CDFS only	2044	17.4	67.3	57.1	58.7	53.2	40.2	54.6
zCOSMOS only	9930	**48.5**	79.7	**74.9**	**75.3**	65.8	60.3	63.2
DEEP2 only	6919	43.7	73.8	17.8	5.5	35.2	**68.8**	**89.5**
G15DEEP only	1792	10.1	42.0	63.1	63.6	44.1	19.8	14.0
VVDS only	4688	34.7	**81.4**	54.9	72.8	**70.7**	57.3	70.2
without CDFS	23329	89.1	98.9	81.5	82.6	82.2	81.0	93.0
without zCOSMOS	15443	77.6	97.8	**76.4**	80.0	80.8	78.9	92.6
without DEEP2	18454	**73.4**	**93.4**	81.6	83.0	80.4	**72.2**	**80.8**
without G15DEEP	23581	90.6	99.5	80.1	83.6	84.3	82.7	94.0
without VVDS	20685	84.7	98.1	81.3	**79.5**	**77.5**	80.1	92.9

The 2D-cosmic shear sample is very well represented by the compilation.

Tomographic binning reduces the representation as noise begins to play a more significant role.

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TOMOGRAPHIC BINNING

Samp/ Photom	All Bins	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
All Spectra	99.5 %	83 %	84 %	85 %	83 %	94 %
Without DEEP2	93 %	82 %	83 %	80 %	72 %	81 %
Without zCOSMOS	98 %	76 %	80 %	81 %	79 %	93 %

The 2D-cosmic shear sample is very well represented by the compilation.

Tomographic binning reduces the representation as noise begins to play a more significant role.

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TOMOGRAPHIC BINNING

Samp/ Photom	All Bins	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
All Spectra	$\boxed{99.5\ \%}$	83 %	84 %	85 %	83 %	94 %
Without DEEP2	93 %	82 %	83 %	80 %	72 %	81 %
Without zCOSMOS	98 %	76 %	80 %	81 %	79 %	93 %

The 2D-cosmic shear sample is very well represented by the compilation.

Tomographic binning reduces the representation as noise begins to play a more significant role.

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TOMOGRAPHIC BINNING

Samp/ Photom	All Bins	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
All Spectra	99.5 %	83 %	84 %	85 %	83 %	94 %
Without DEEP2	93 %	82 %	83 %	80 %	72 %	81 %
Without zCOSMOS	98 %	76 %	80 %	81 %	79 %	93 %

Our highest-z tomographic bin is our *best* represented. DEEP2 informs best the highest redshift bins. zCOSMOS the lowest.

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TOMOGRAPHIC BINNING

Samp/ Photom	All Bins	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
All Spectra	99.5 %	83 %	84 %	85 %	83 %	94 %
Without DEEP2	93 %	82 %	83 %	80 %	72 %	81 %
Without zCOSMOS	98 %	76 %	80 %	81 %	79 %	93 %

Our highest-z tomographic bin is our *best* represented.

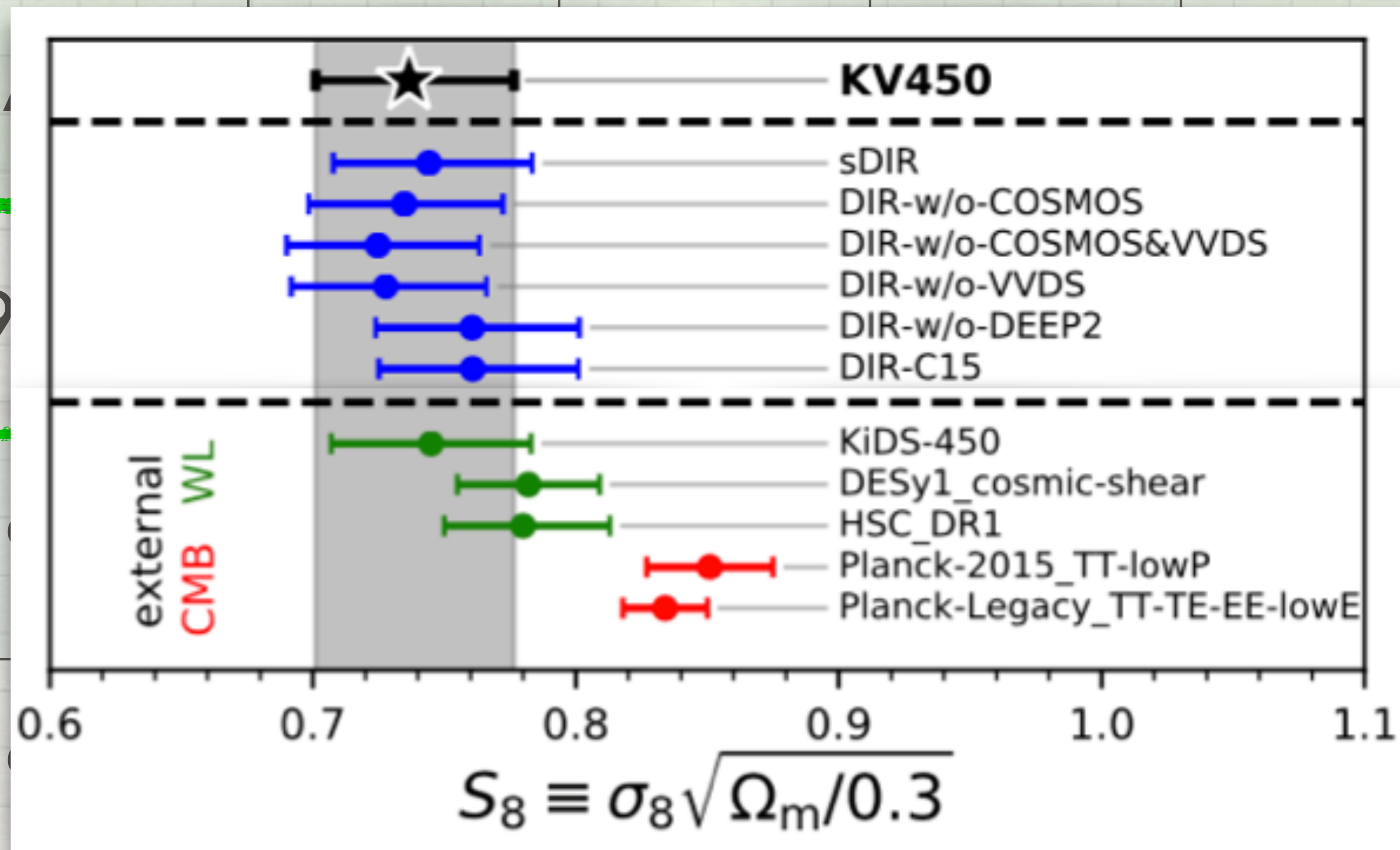
DEEP2 informs best the highest redshift bins. zCOSMOS the lowest.

Wright et al (submitted)

SPECTROSCOPIC VALIDATION FOR COSMIC SHEAR

TOMOGRAPHIC BINNING

Samp/ Photom	Bin 5
All Spectra	94 %
Without DEEP2	81 %
Without zCOSMOS	93 %

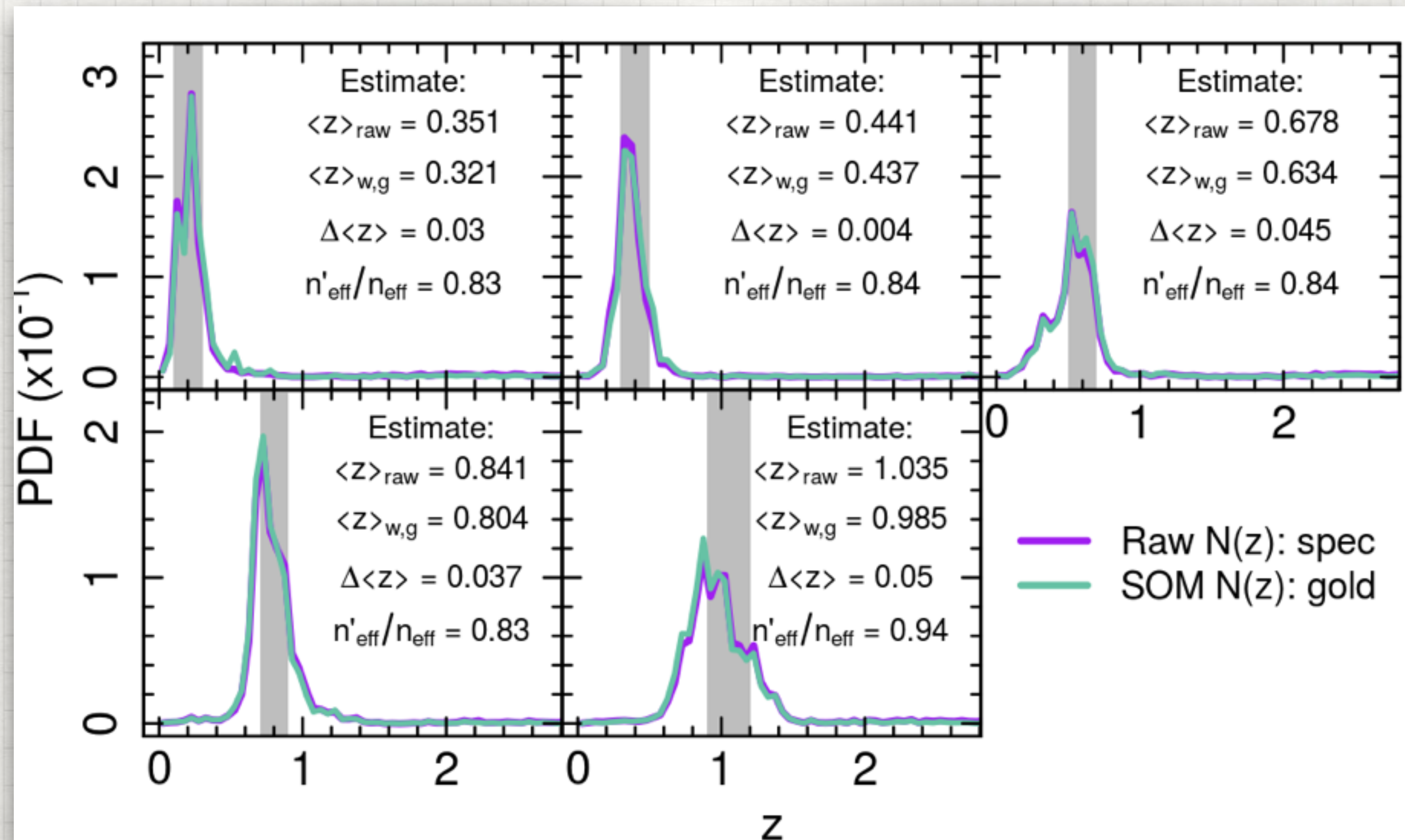


Our highest-z tomographic bin is our *best* represented. DEEP2 informs best the highest redshift bins. zCOSMOS the lowest.

Wright et al (submitted)

NEW PHOTOMETRIC REDSHIFT CALIBRATION

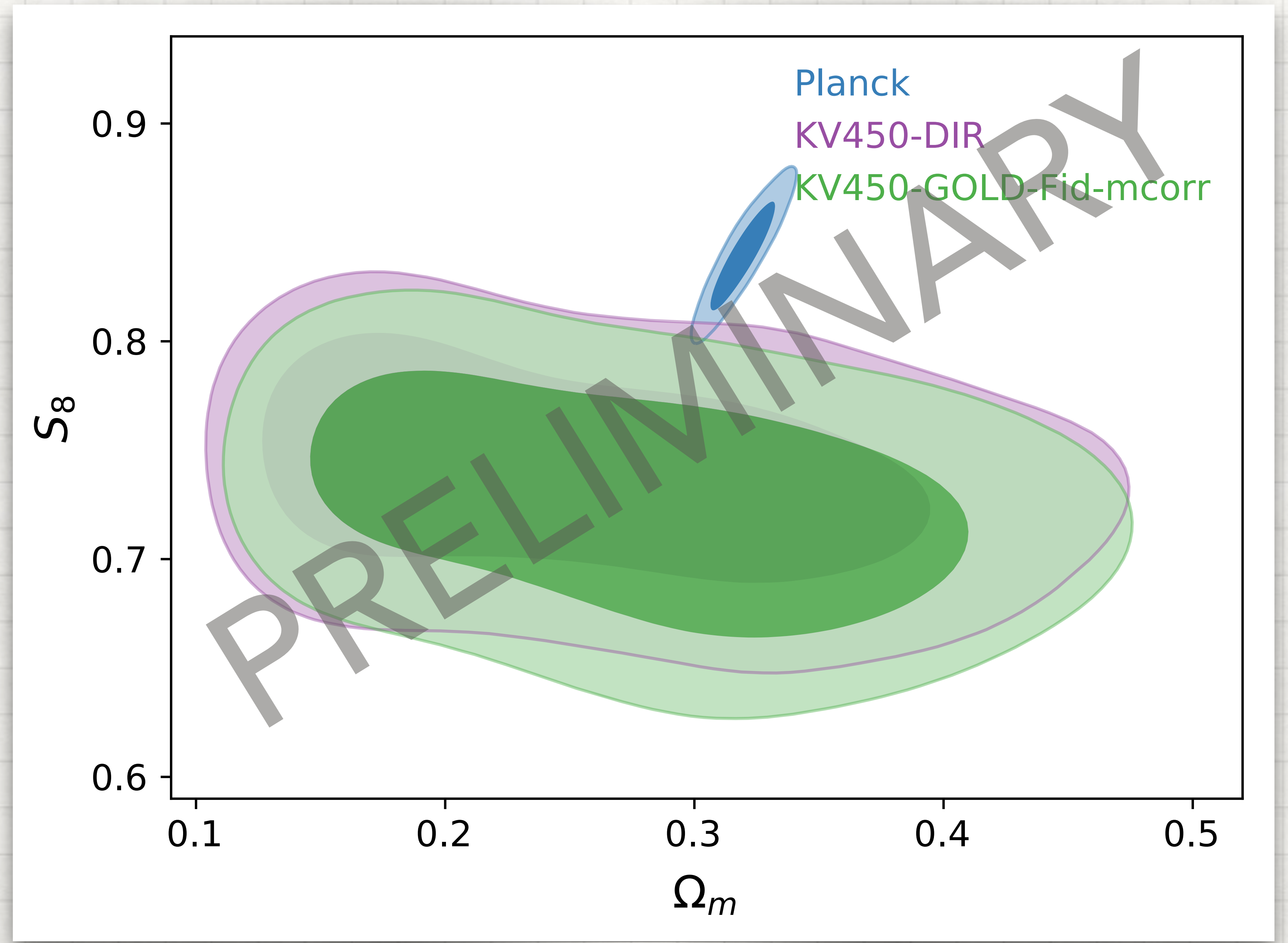
KV450 DATA



Wright et al (submitted)

NEW COSMOLOGICAL PARAMETER ESTIMATES

KV450 GOLD SAMPLE



Wright et al (in prep)

SUMMARY

- At KiDS-level noise, the MICE2 redshift distributions are unbiased in all bins except for bin 4, but this is limited to MICE2.
- In the noiseless case, the SOM redshift calibration is unbiased.
- The bin 4 bias is caused by a colour-redshift degeneracy present in the DEEP2 colour-space.
  Future analyses need to be robust to this.
- The dominant factor in the accuracy of the redshift calibration is noise in the spectroscopic compilation.
- The SOM calibration method allows construction of a 100% represented 'gold' sample, which we will use for a reanalysis of KV450 cosmology.
- The KiDS spectroscopic compilation represents >99% of the 2D photometric lens sample.