
46
FIAT
YAMAHA
FASTWEB
MOTUL
YAMAHA
motogp.com
official website

Estimating Photometric Redshifts using Gaussian Process Regression in the Sloan Digital Sky Survey

(and 2MASS)

Michael Way (NASA/Goddard Institute for Space Studies)
Paul Gazis, Jeffrey Scargle (NASA/Ames, Space Sciences Division)
Ashok Srivastava (NASA/Ames Intelligent Systems Division)
Les Foster + Students (San Jose State University)

http://astrophysics.arc.nasa.gov/~mway/Stanford-200807.pdf

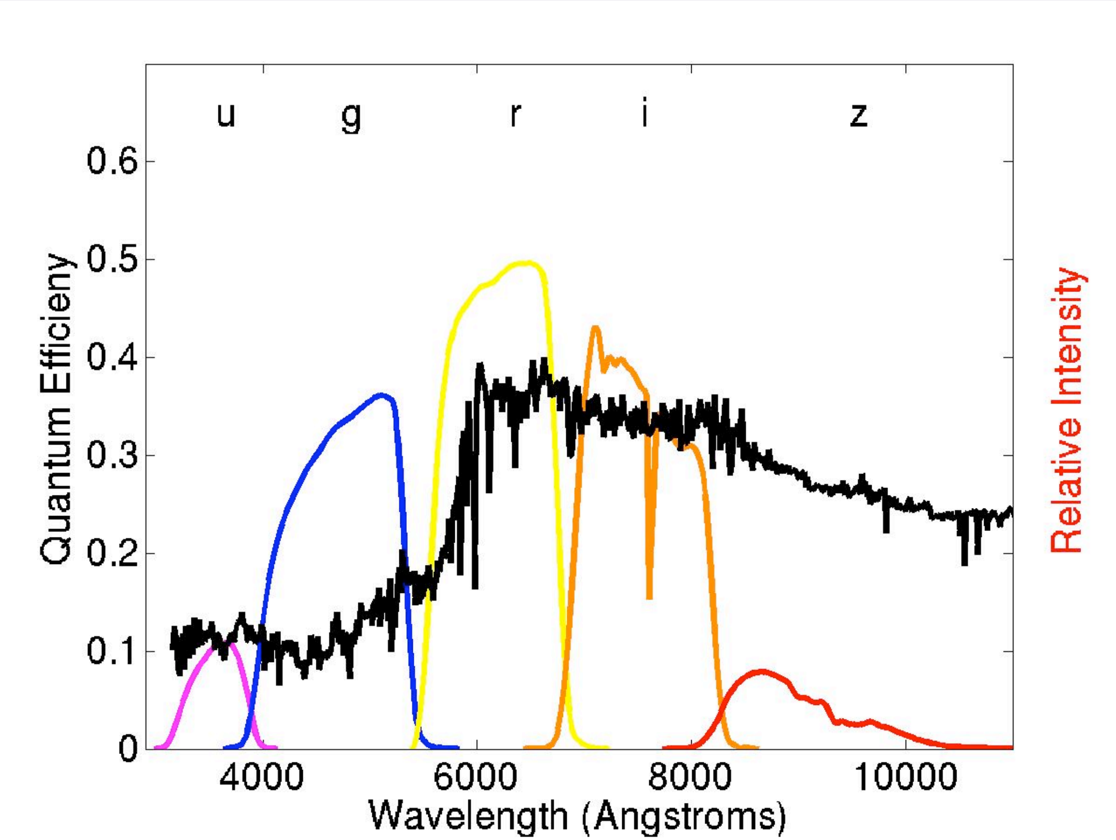
Outline

- What are Photometric Redshifts?
- A quick background on common training set methods
- What is Gaussian Process Regression?
- Do different kinds of Kernels matter?
- Does better quality photometry matter?
- How many galaxies do I need to get a good fit?
- Do SDSS morphological indicators help?

What are Photometric Redshifts?

Photometric Redshifts: A **rough** estimate of the redshift of a galaxy without having to measure a spectrum.

$$Z_{spec}=(\lambda_{measured}-\lambda_{rest})/\lambda_{rest} \qquad z_{photo}=z(C,m)$$

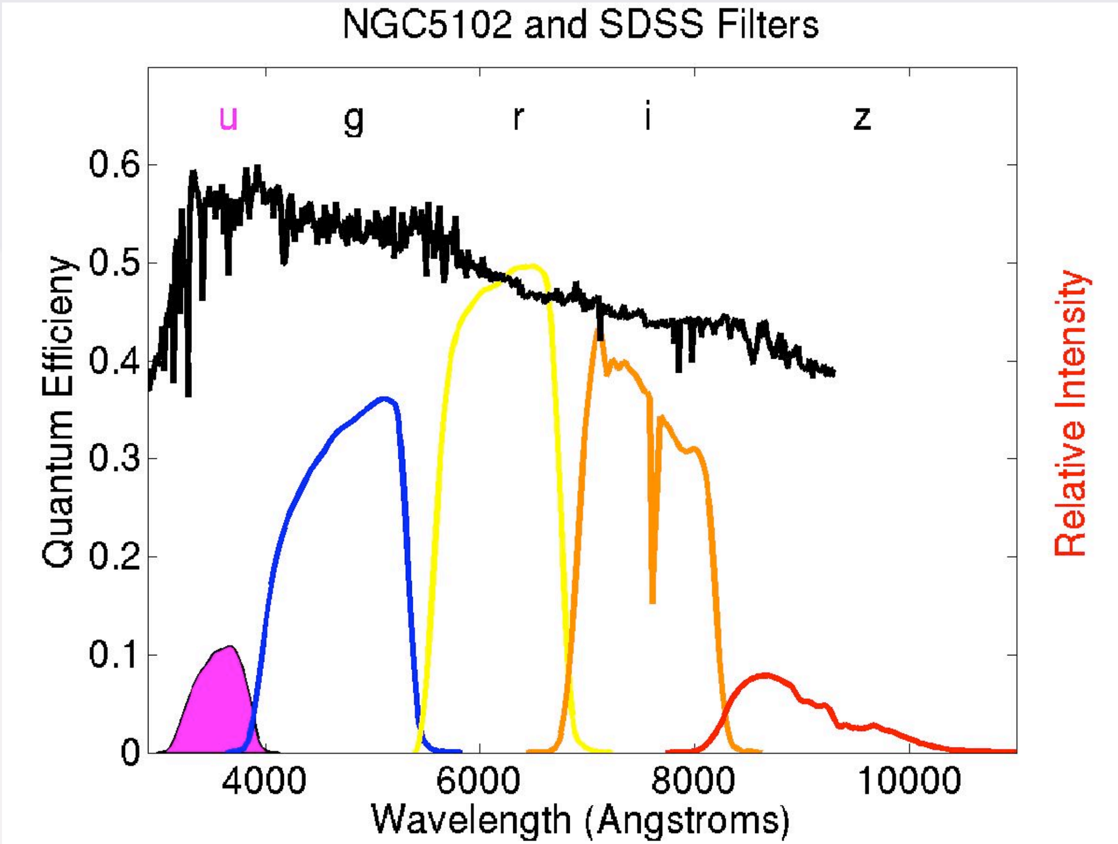


What are Photometric Redshifts?

$$\mathbf{Z_{spec}=(\lambda_{measured}-\lambda_{rest})/\lambda_{rest}}$$

$z_{photo}=z(C,m)$

z=0.0

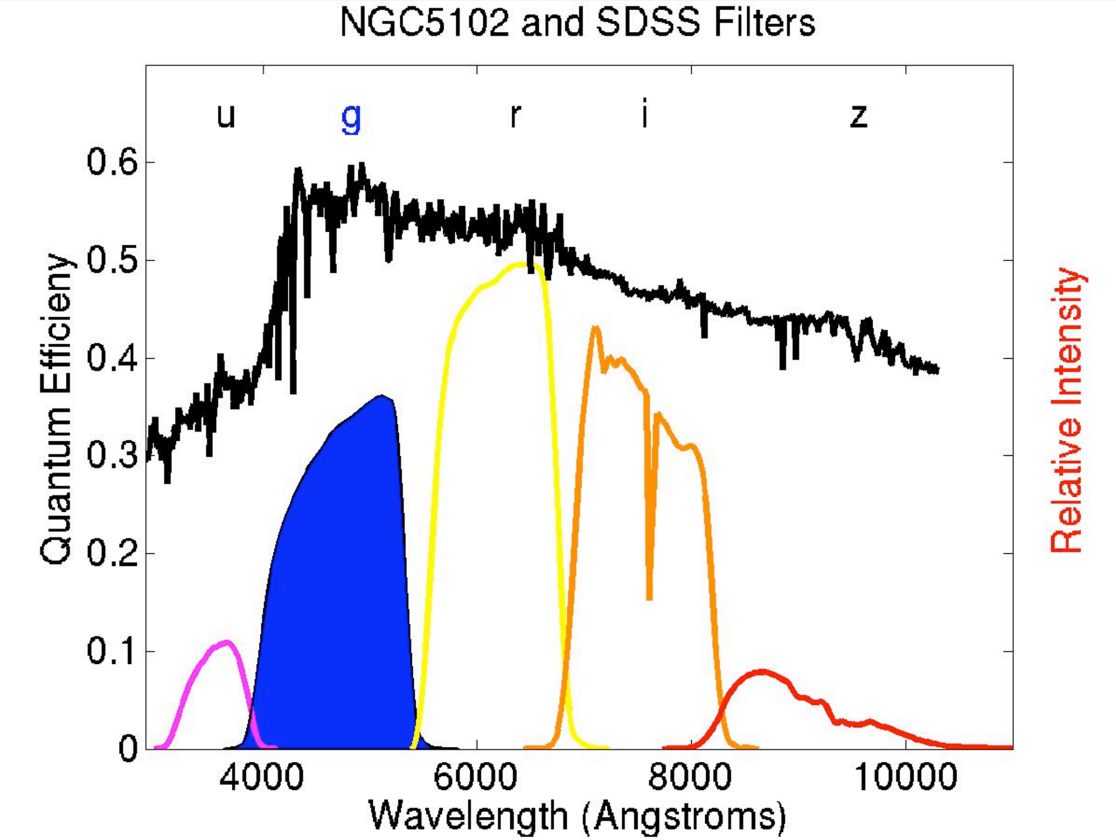


Stanford 08

What are Photometric Redshifts?

$$z_{spec} = (\lambda_{measured} - \lambda_{rest})/\lambda_{rest} \qquad z_{photo} = z(C,m)$$

z~0.06 (18000 km/s)



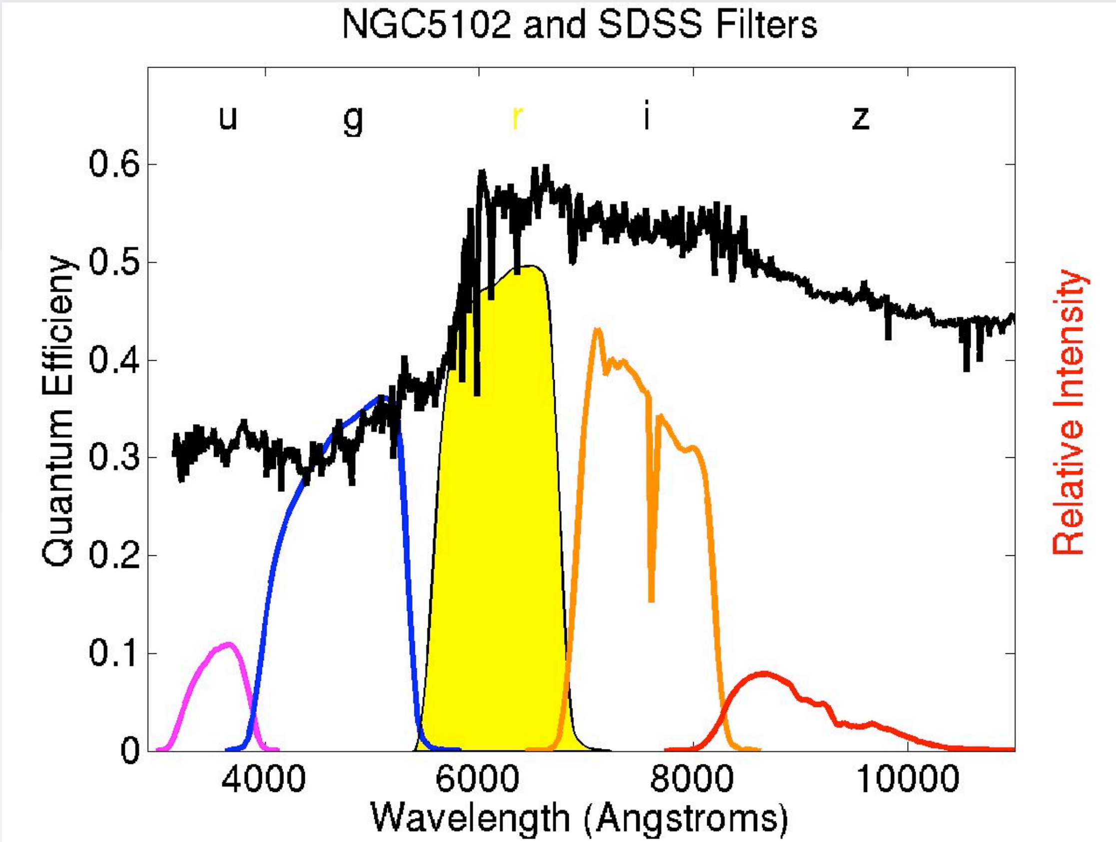
Stanford 08

What are Photometric Redshifts?

$$\mathbf{z_{spec}=(\lambda_{measured}-\lambda_{rest})/\lambda_{rest}}$$

$$\mathbf{z \sim 0.6}$$

$$z_{photo}=z(C,m)$$



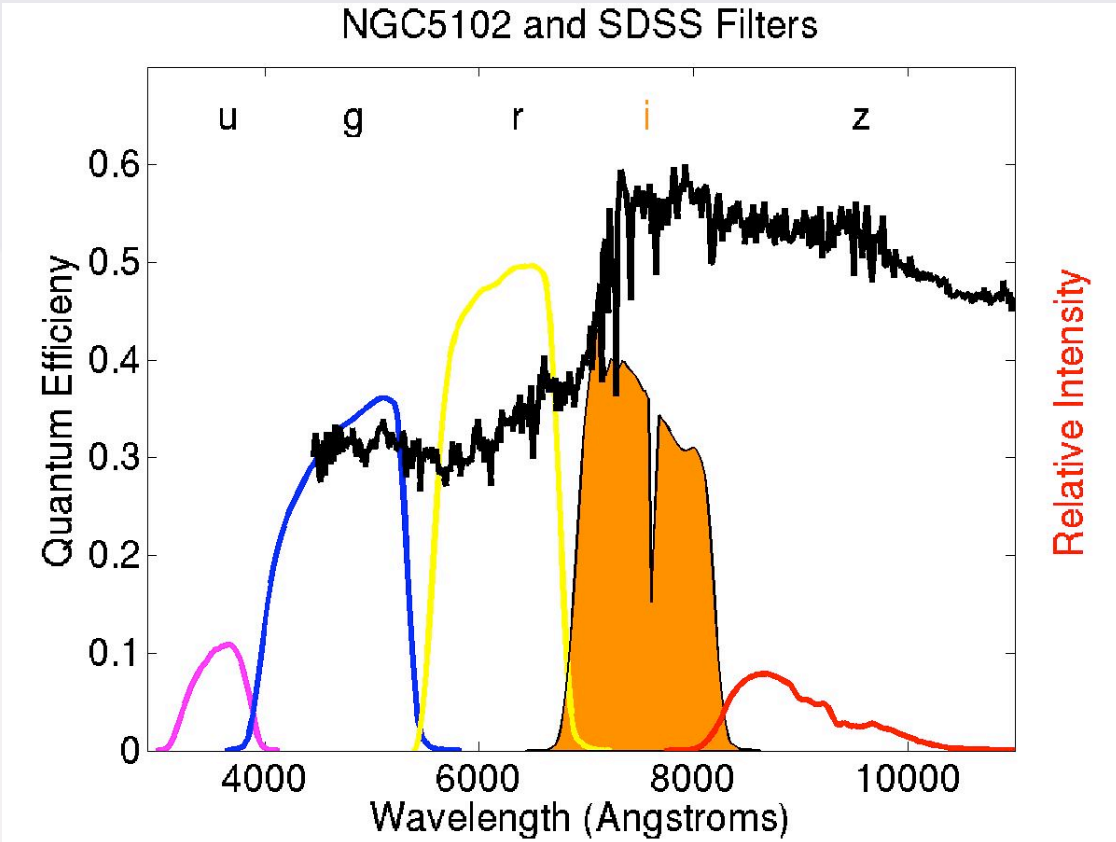
Stanford 08

What are Photometric Redshifts?

$$Z_{spec} = (\lambda_{measured} - \lambda_{rest})/\lambda_{rest}$$

$$z_{photo} = z(C, m)$$

z~0.90

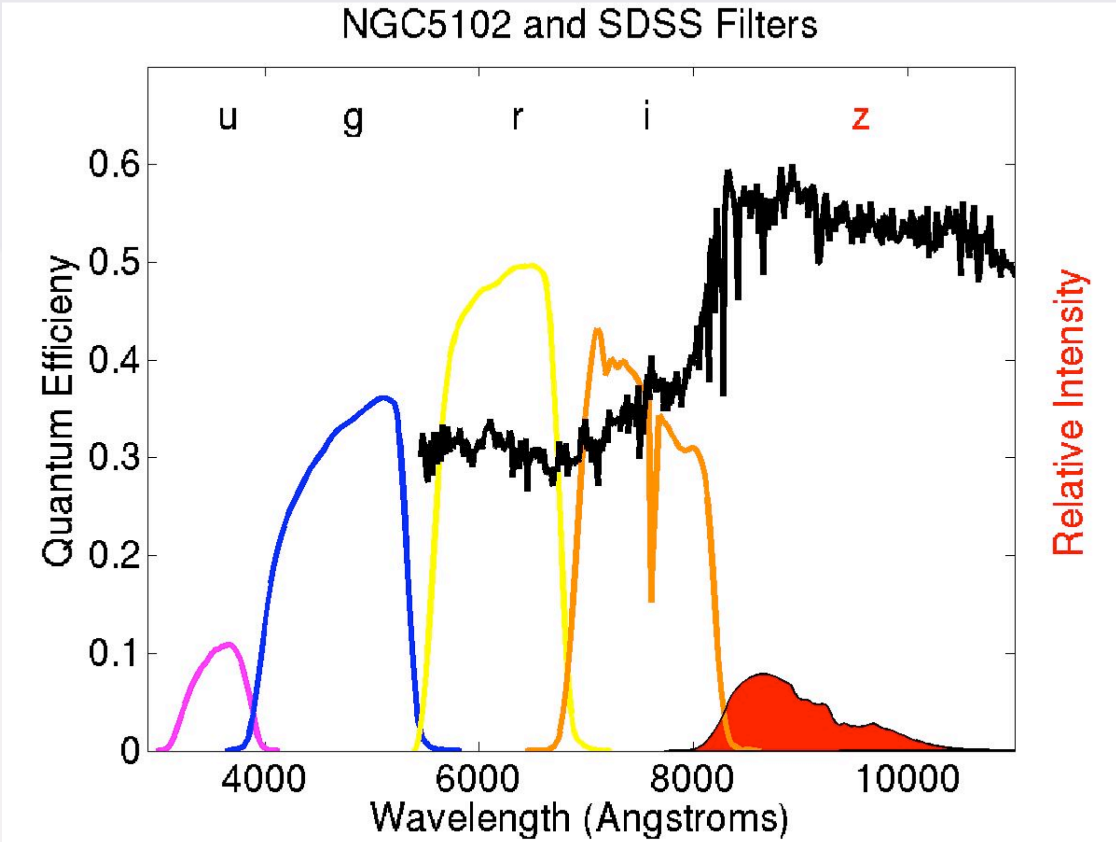


Stanford 08

What are Photometric Redshifts?

$Z_{spec}=(\lambda_{measured}-\lambda_{rest})/\lambda_{rest}$ $z_{photo}=z(C,m)$

z~1.10



Stanford 08

Photo-z methods

1.) Spectral Energy Distribution (SED) Fitting:

- model based approach
- uses redshifts derived from spectra of artificial galaxies (e.g. Bruzual & Charlot)

2.) Training-Set methods:

- **empirical approach**
- **uses *spectroscopic* redshifts from a sub-sample of galaxies with the same band-pass filters**

Photo-z The Empirical Approach

Training Set Methods need a sub-sample of Galaxies:

- of known spectroscopic redshift

- with a comparable range of **magnitudes** **(u g r i z)** to our Photometric survey objects

- These will be our "Training Samples"

“Training Set” Methods

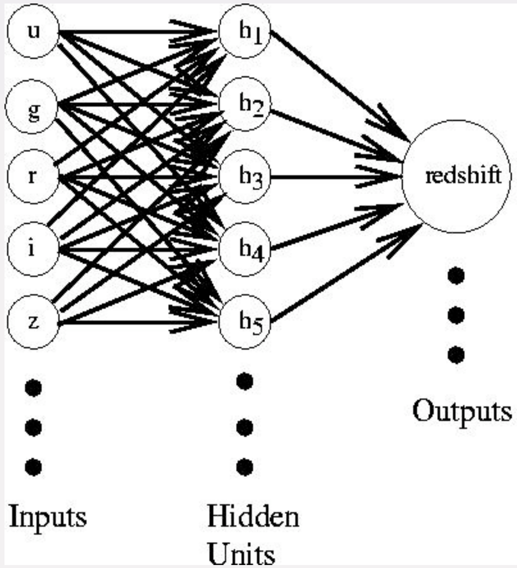
Galaxy Photometric Redshift Prediction History

- Linear Regression was first tried in the 1960s
- Quadratic & Cubic Regression (1970s)
- Polynomial Regression (1980s)
- Neural Networks (1990s)
- Kd Trees & Bayesian Classification Approaches (1990s)
- Support Vector Machines & GP Regression (2000s)

Gaussian Process Regression ⇔ Kernel Methods

Kernel Methods have replaced Neural Networks in the Machine Learning literature

WHY?: given a large # of hidden units => GP (Neal 1996).



$h_n > 100$

→→→→→

Stanford 08

Kernel Methods - Gaussian Process Regression

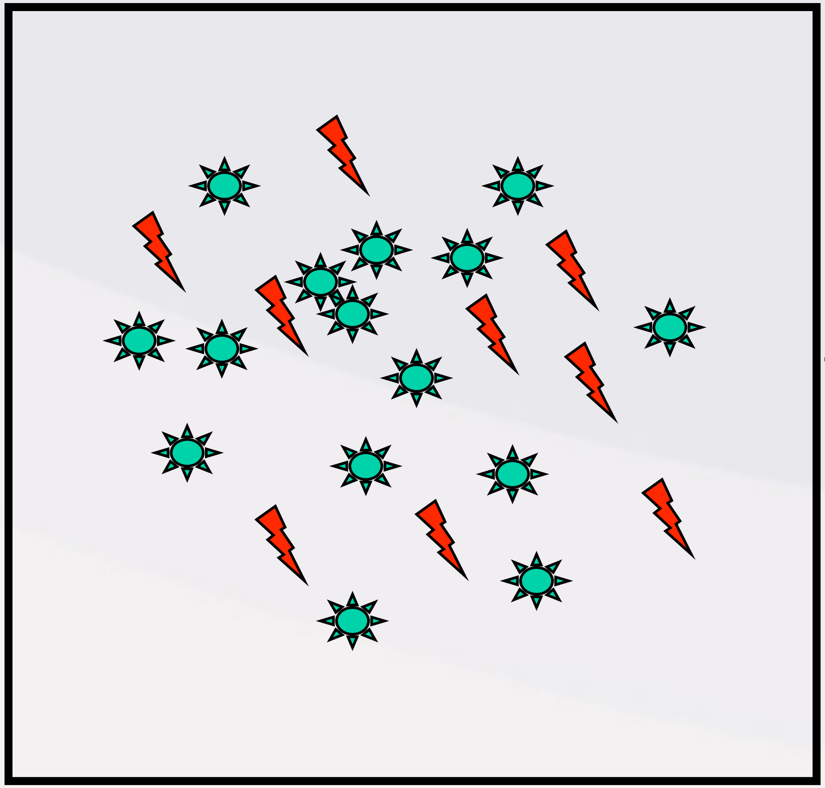
GP regression builds a linear model in a very high dimensional *parameter space* ("feature space" → Hilbert space).

- One can map the data using a function F(x) [kernel] into this high (or infinite) dimensional *parameter space* where one can perform linear operations.

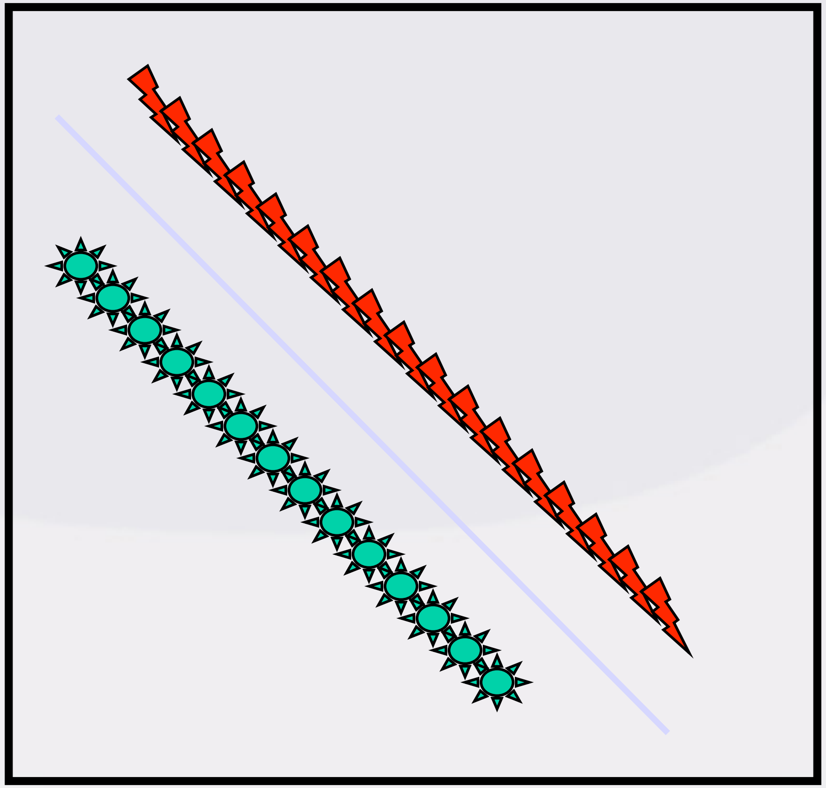
The value of kernels

Original Data without Kernel

Mapped Data using Kernel



F(x)
Kernel
Map



Data in original space: highly complex decision boundaries.

Data in high dimensional feature space after mapping through F(x) can yield simple decision boundaries.

GP Regression (Kernels)

GP Advantages:

- Small input data training samples (good for higher redshifts?) yet low errors

- Over fitting is eliminated by use of proper priors

- Realistic estimation of individual redshift errors

GP Disadvantages:

- Possibly large CPU time requirements
  - The Kernel (Covariance Matrix) **can** be large: $K=(\lambda^2 I+XX^T)^2$ if $X=5x180{,}000$ (our case) then K is a matrix 180,000 x 180,000 and we have:

    $$y^* = K^*(\lambda^2 I + K)^{-1} y$$

  - Need to invert this large K matrix - $O(N^3)$ operation
- Kernel Selection is ambiguous?

GP: Which Kernel??

Using GPs Part I: Pick a transfer/covariance function

Matern Class Fcn

$$k(r) = \frac{2^{l-v}}{\Gamma(v)}\left(\frac{\sqrt{2vr}}{l}\right)^{v} J_v\left(\frac{\sqrt{2vr}}{l}\right)$$

$v \rightarrow \infty$

Radial Basis Fcn

$$k(r) = \exp\left(\frac{r^2}{2l^2}\right)$$

Rational Quadratic

$$k_{RQ}(r) = 1 + \left(\frac{r^2}{2\alpha l^2}\right)^{-\alpha}$$

Polynomial

$$k(x,x') = \left(\sigma_o^2 + x^T \sum_p x'\right)^p$$

Neural Nets

$$k_{NN}(x,x') = \frac{2}{\pi}\sin^{-1}\left(\frac{2x^T\Sigma x'}{\sqrt{(1+2x^T\Sigma x)(1+2x'^T\Sigma x')}}\right)$$

GP Regression How-to

Using GPs Part II: That matrix inversion…

With our SDSS (DR3) Main Galaxy spectroscopic sample (180,000 galaxies) the matrix size is 180,000 x 180,000

- Need a SSI supercomputer with a LOT of ram and cpu time?
- One can take a random sample of ~1000 galaxies & invert that while bootstrapping n times from full sample (Paper I)
- **However, some *low-rank matrix approximations* work well** (Cholesky Decomposition, Subset of Regressors, Projected Process Approx, etc.)

Results: Other authors

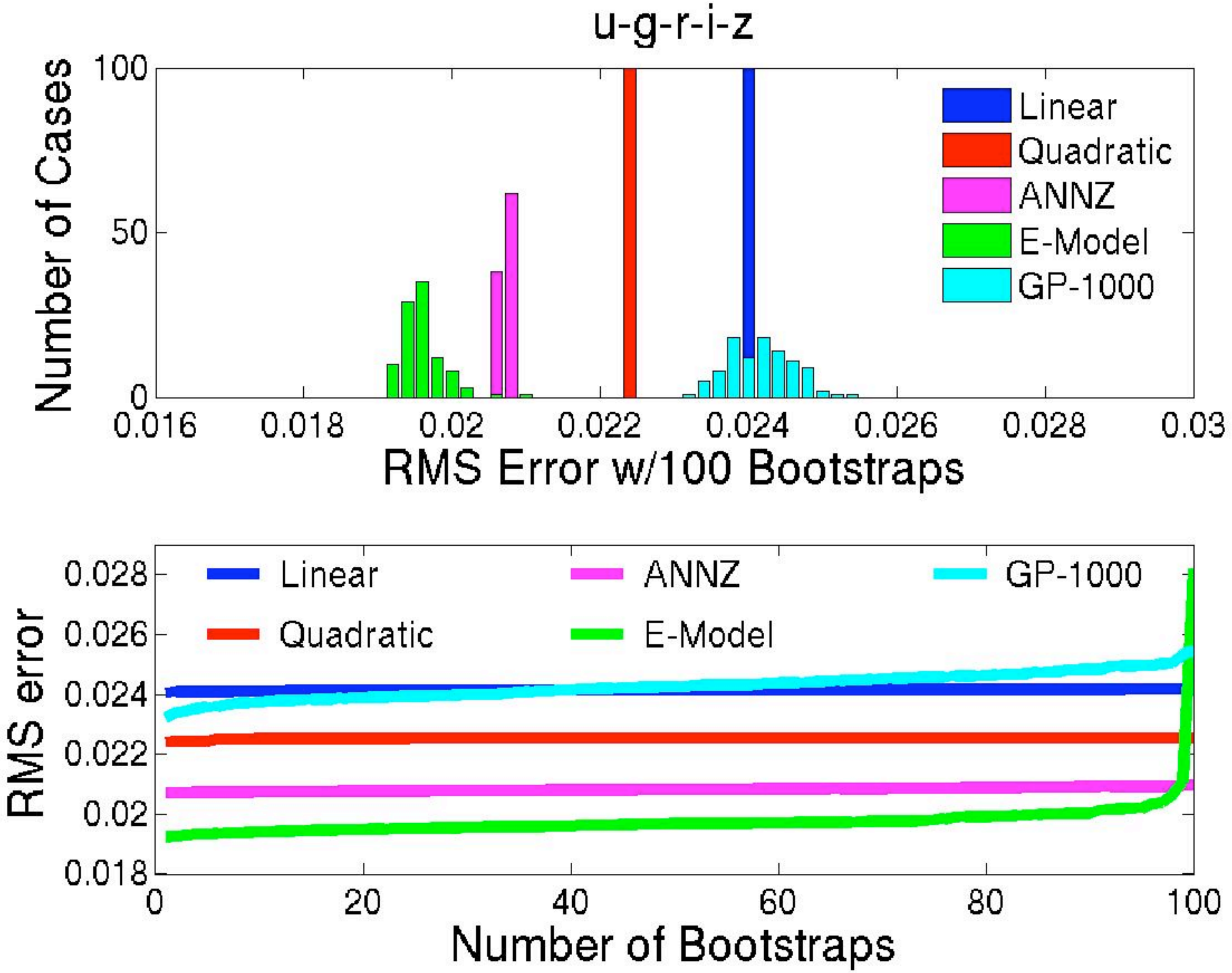
Method Name	σ_{rms}	Dataset[1]	Inputs[2]	Source
CWW	0.0666	SDSS-EDR	ugriz	Csabai et al. (2003)
Bruzual-Charlot	0.0552	SDSS-EDR	ugriz	Csabai et al. (2003)
ClassX	0.0340	SDSS-DR2	ugriz	Suchkov et al. (2005)
Polynomial	0.0318	SDSS-EDR	ugriz	Csabai et al. (2003)
Support Vector Machine	0.0270	SDSS-DR2	ugriz	Wadadekar (2005)
Kd-tree	0.0254	SDSS-EDR	ugriz	Csabai et al. (2003)
Support Vector Machine	0.0230	SDSS-DR2	ugriz+r50+r90	Wadadekar (2005)
Artificial Neural Network	0.0229	SDSS-DR1	ugriz	Collister & Lahav (2004)

GP Regression (Results)

Results: SDSS (DR3) Main Galaxy Sample

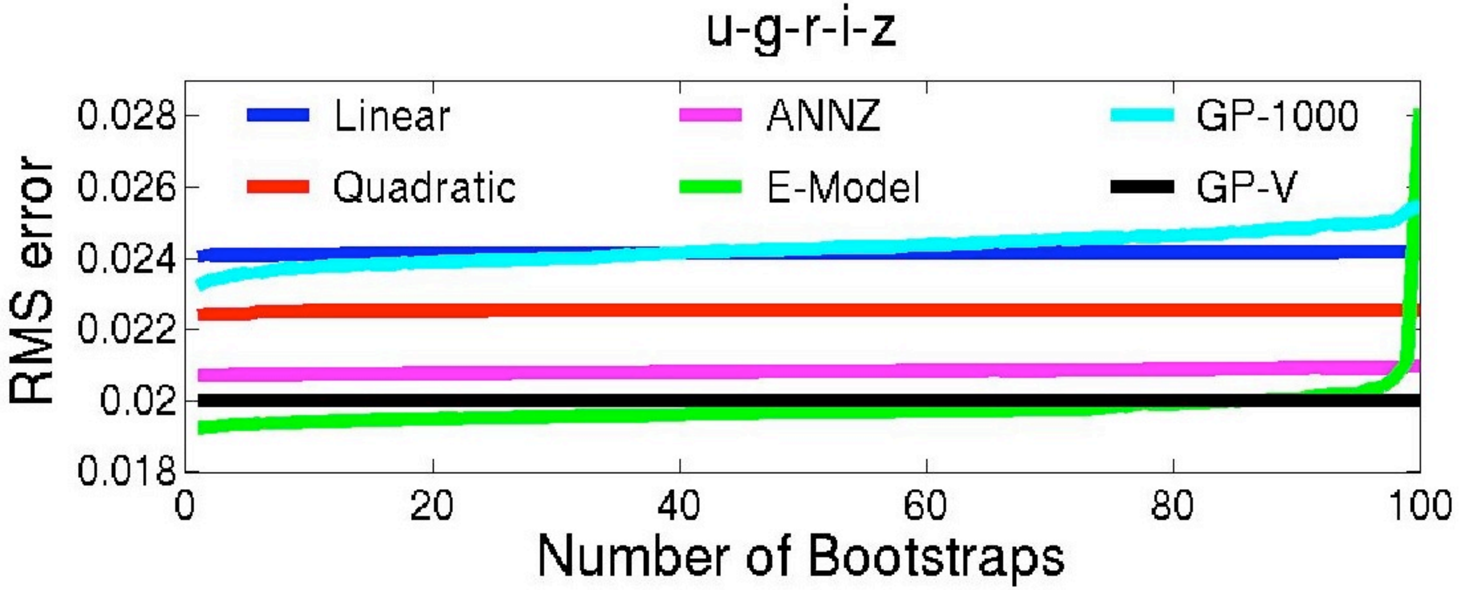
- Paper I: Compared linear, quadratic, Neural Networks and GPs on the SDSS
- With ONLY 1000 samples GPs performed well compared to the other methods
- Paper II: With *low-rank matrix inversion approximations* GPs performed better than all other methods

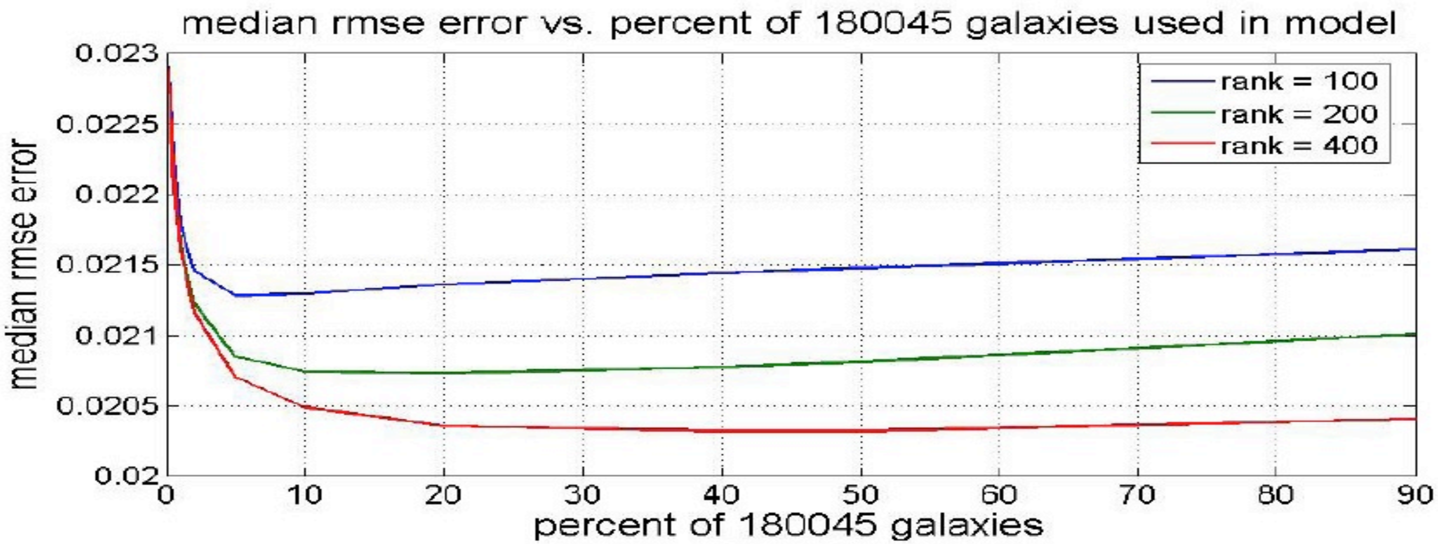
Paper I Results: Comparing Methods



Latest Results: Comparing Methods

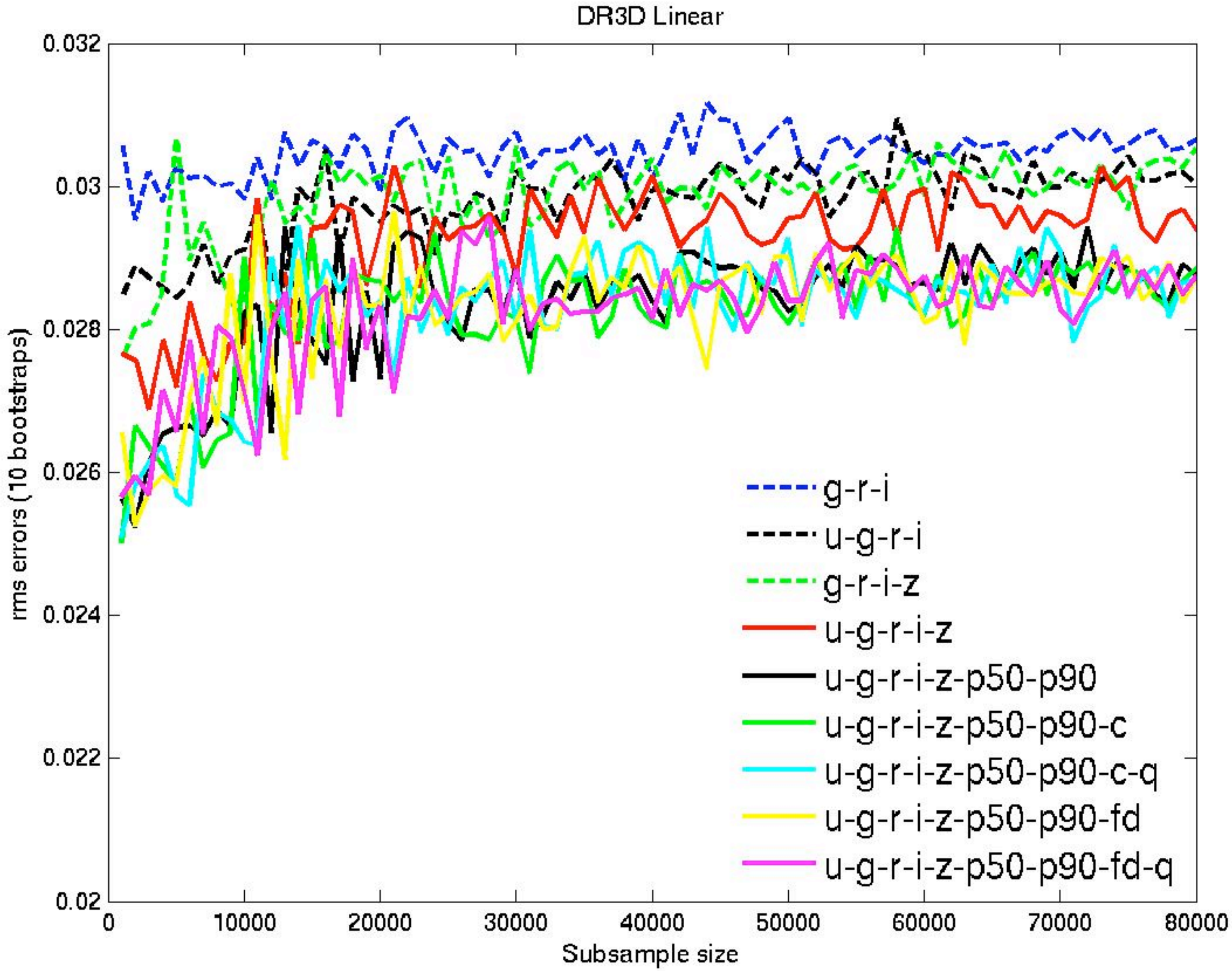
→ GP-V: Rank=1000 for 36000



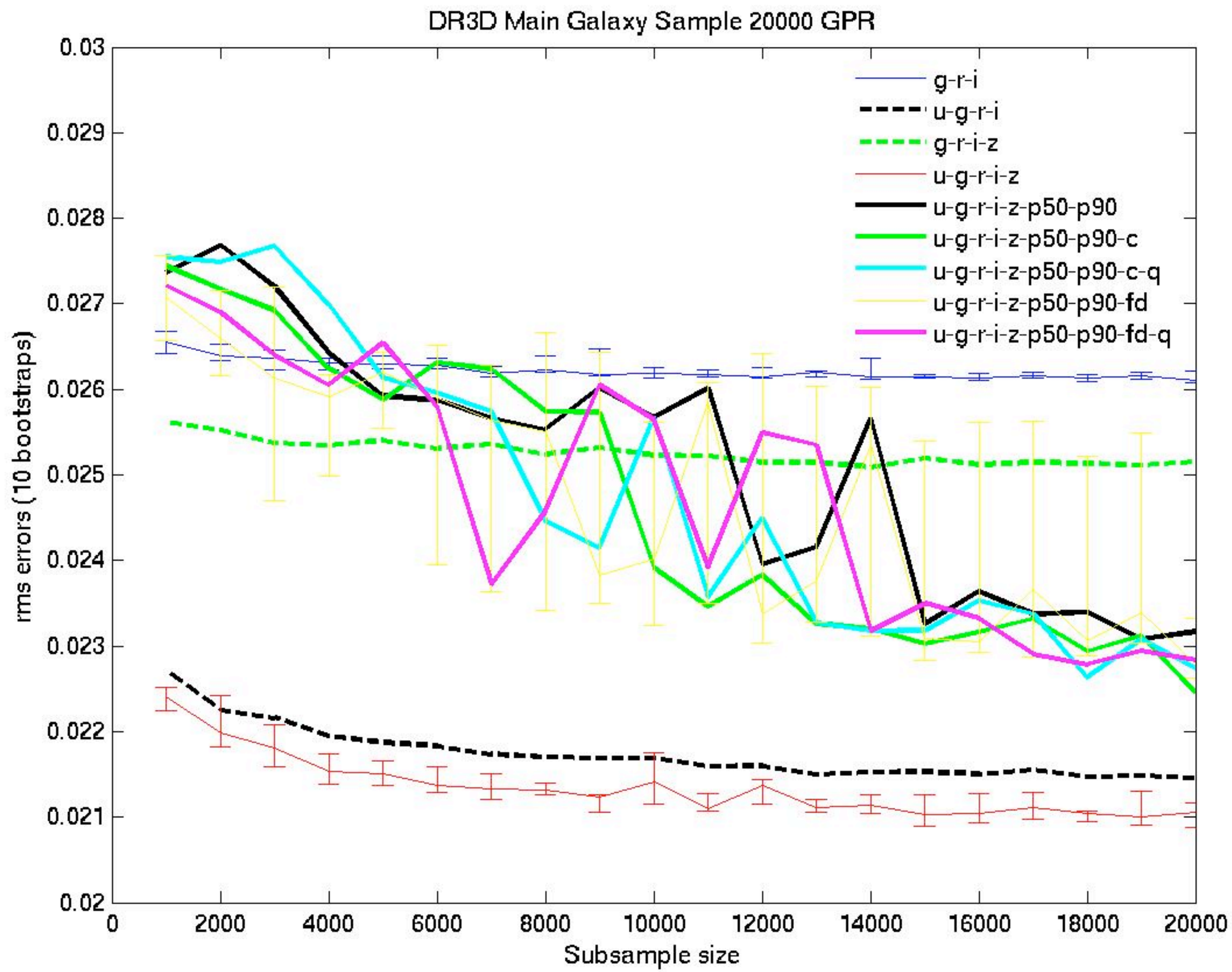


Beyond 20% (~36000) Rank 400 is fairly flat ←

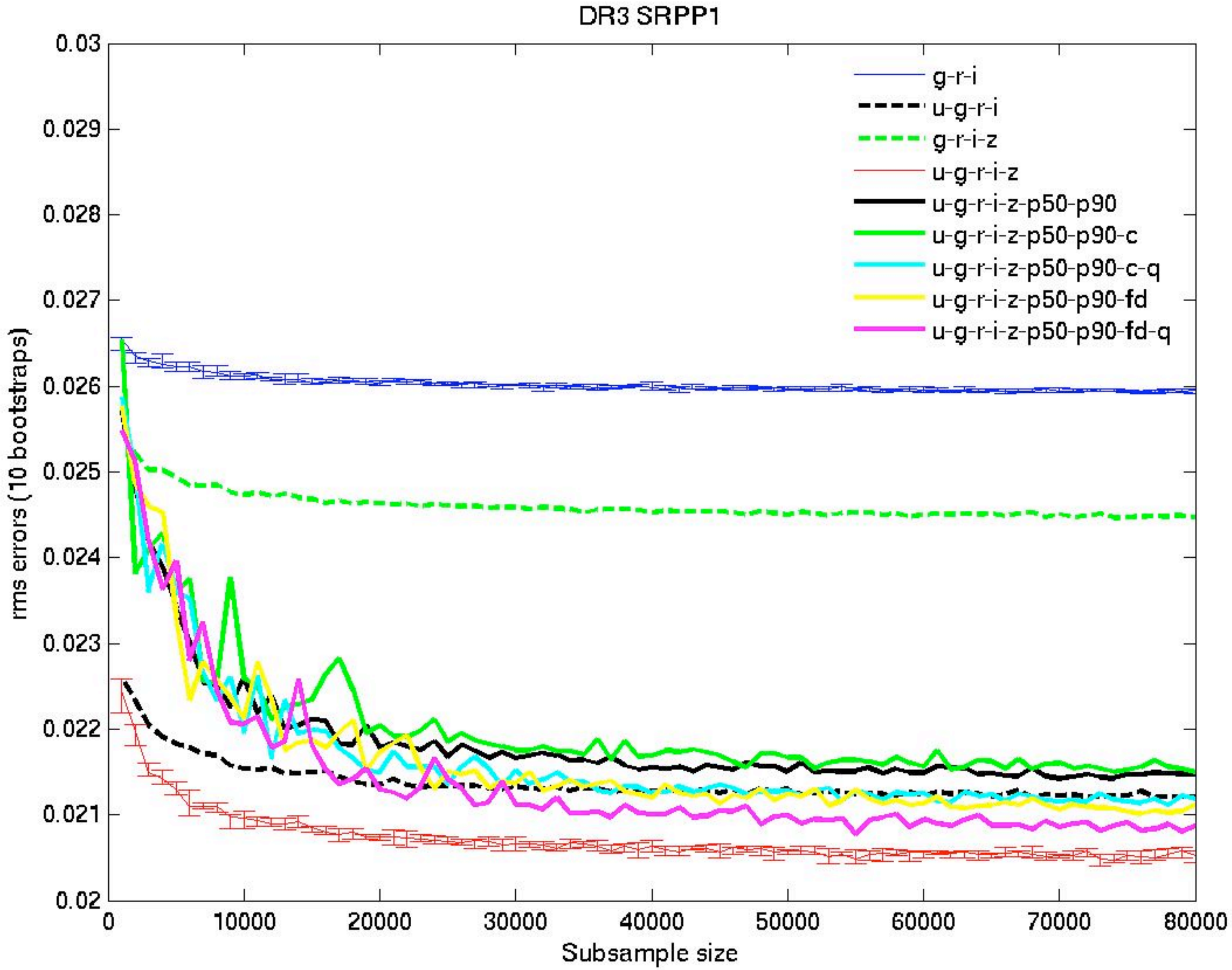
Main Galaxy Sample 80000 Linear



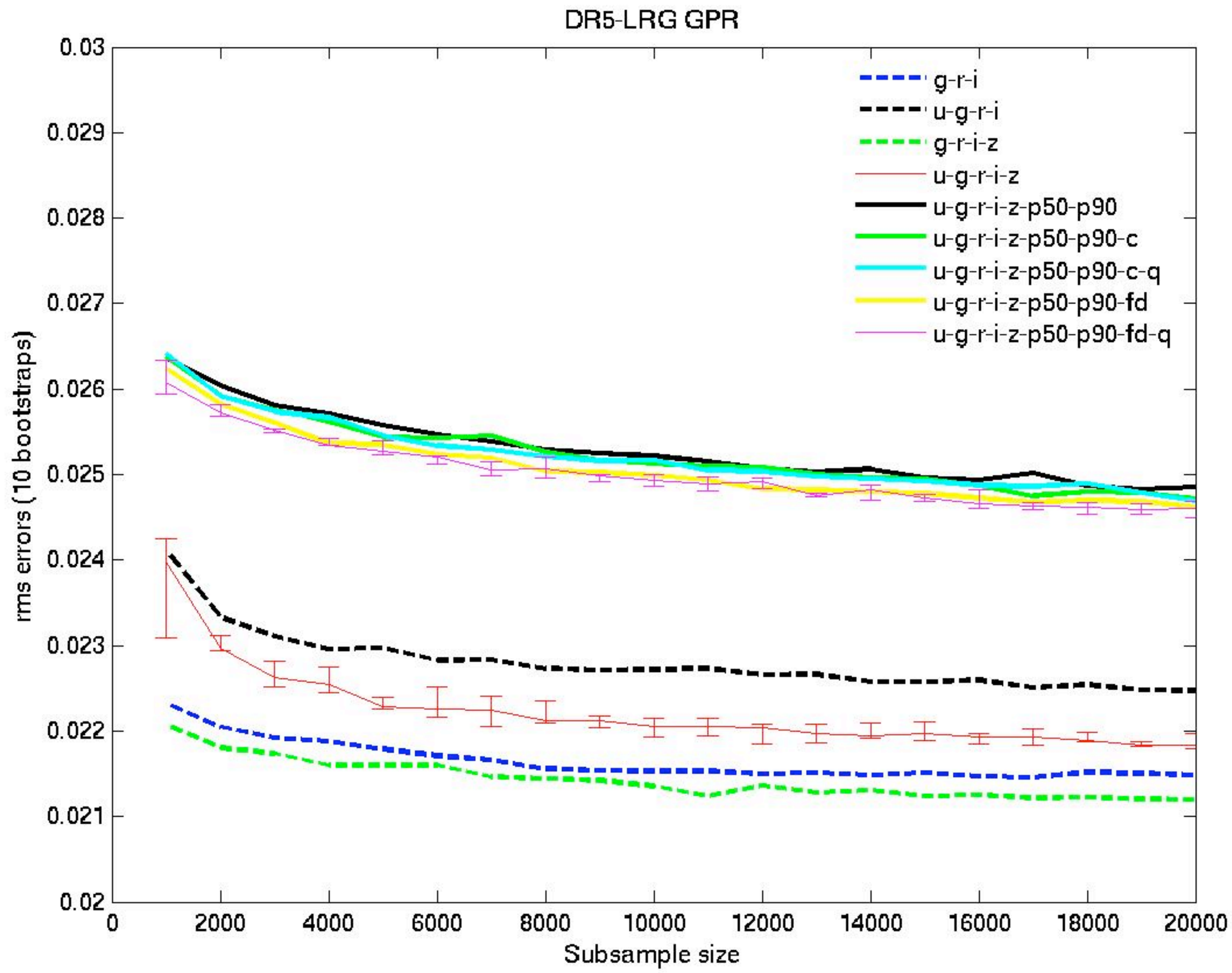
Main Galaxy Sample 20000 GPR



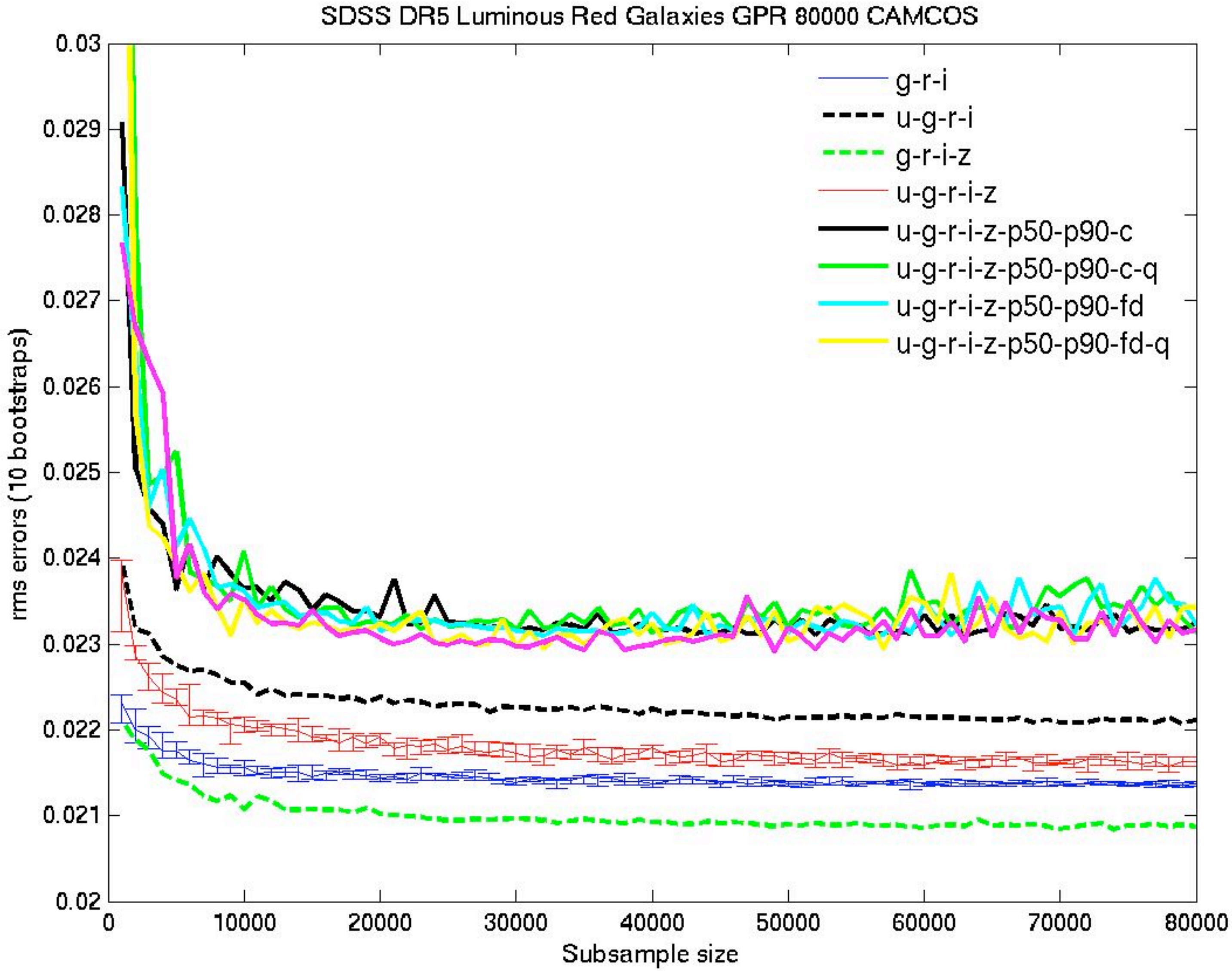
Main Galaxy Sample 80000 GPR



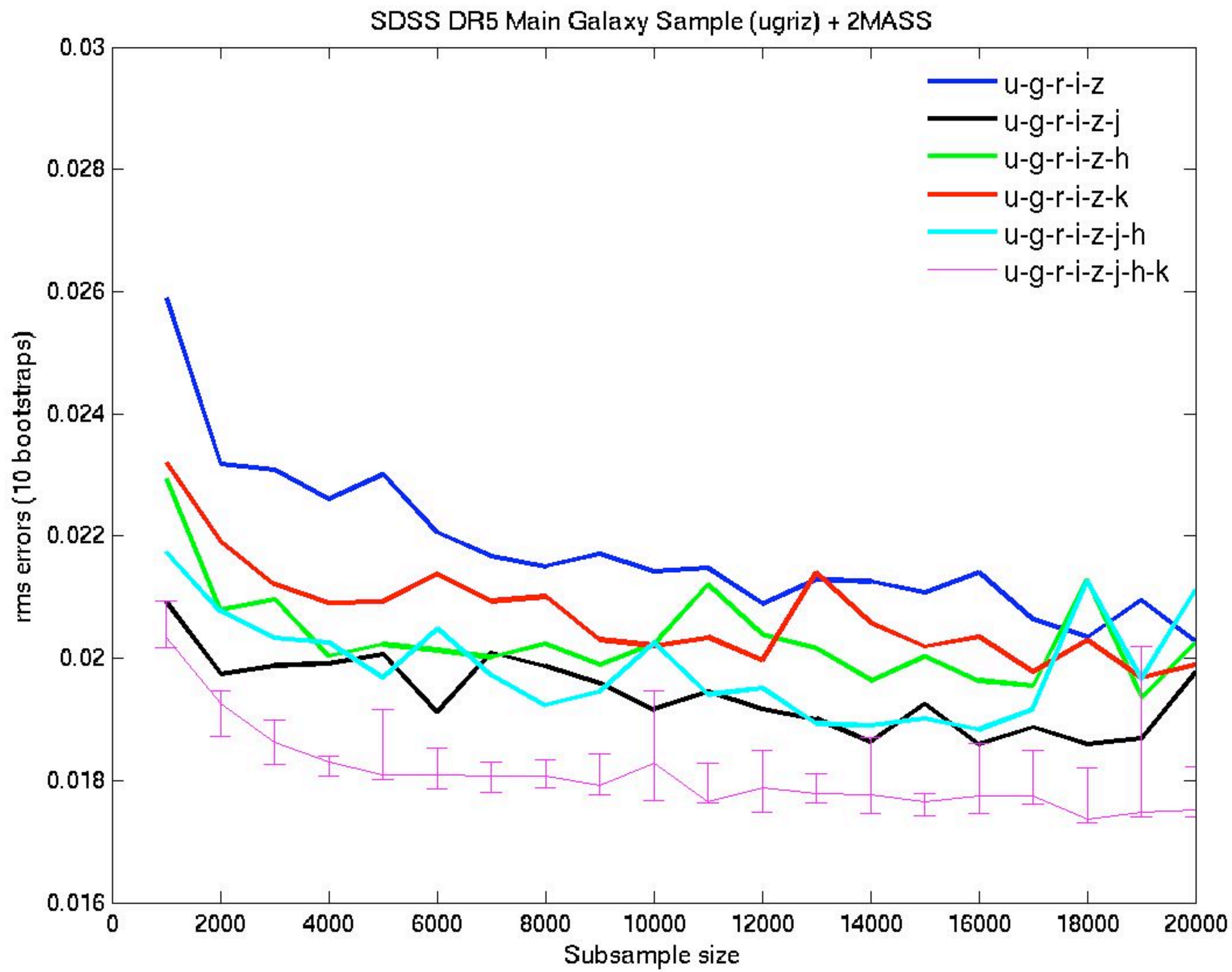
Luminous Red Galaxies 20000 GPR



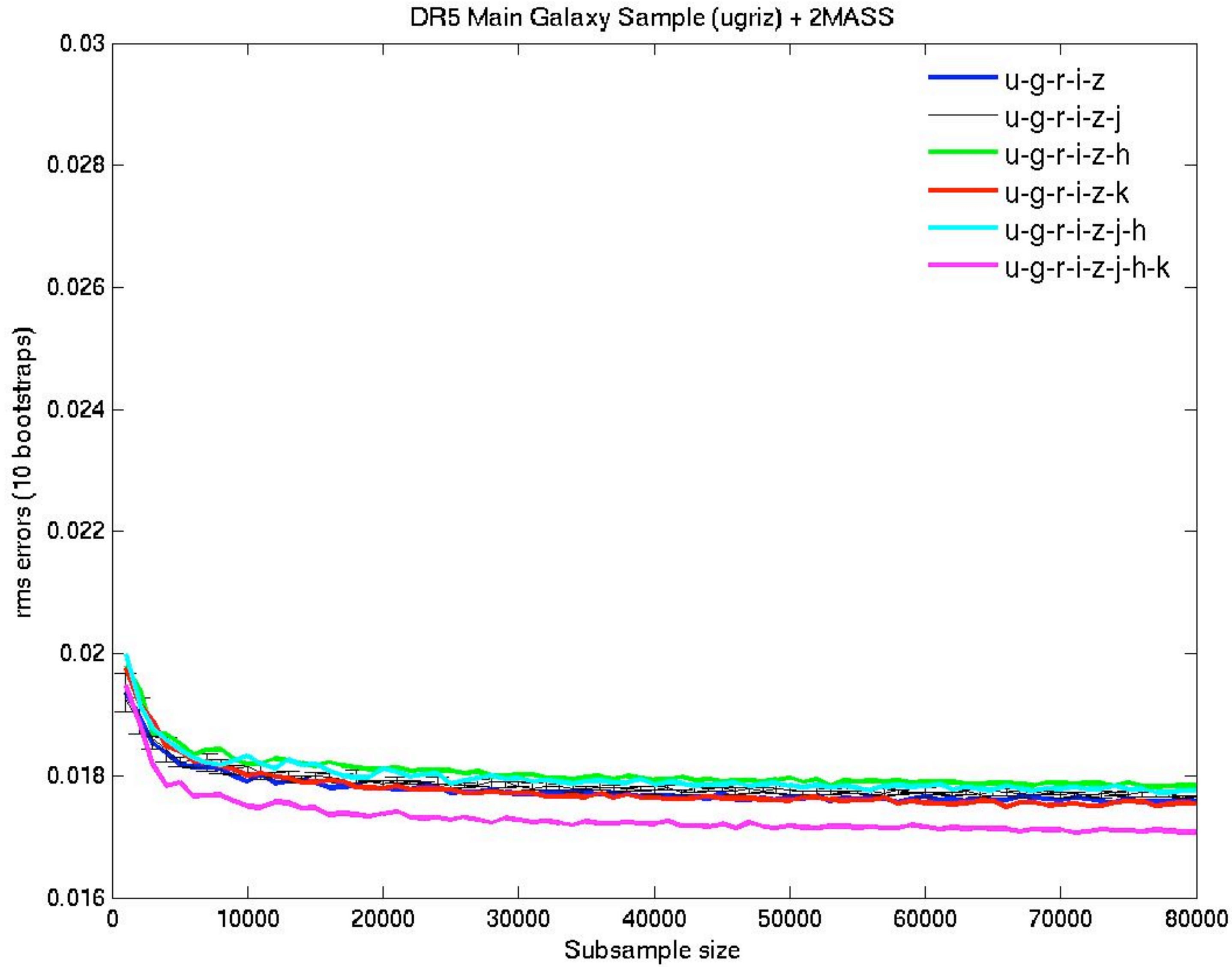
LRG 80000 GPR



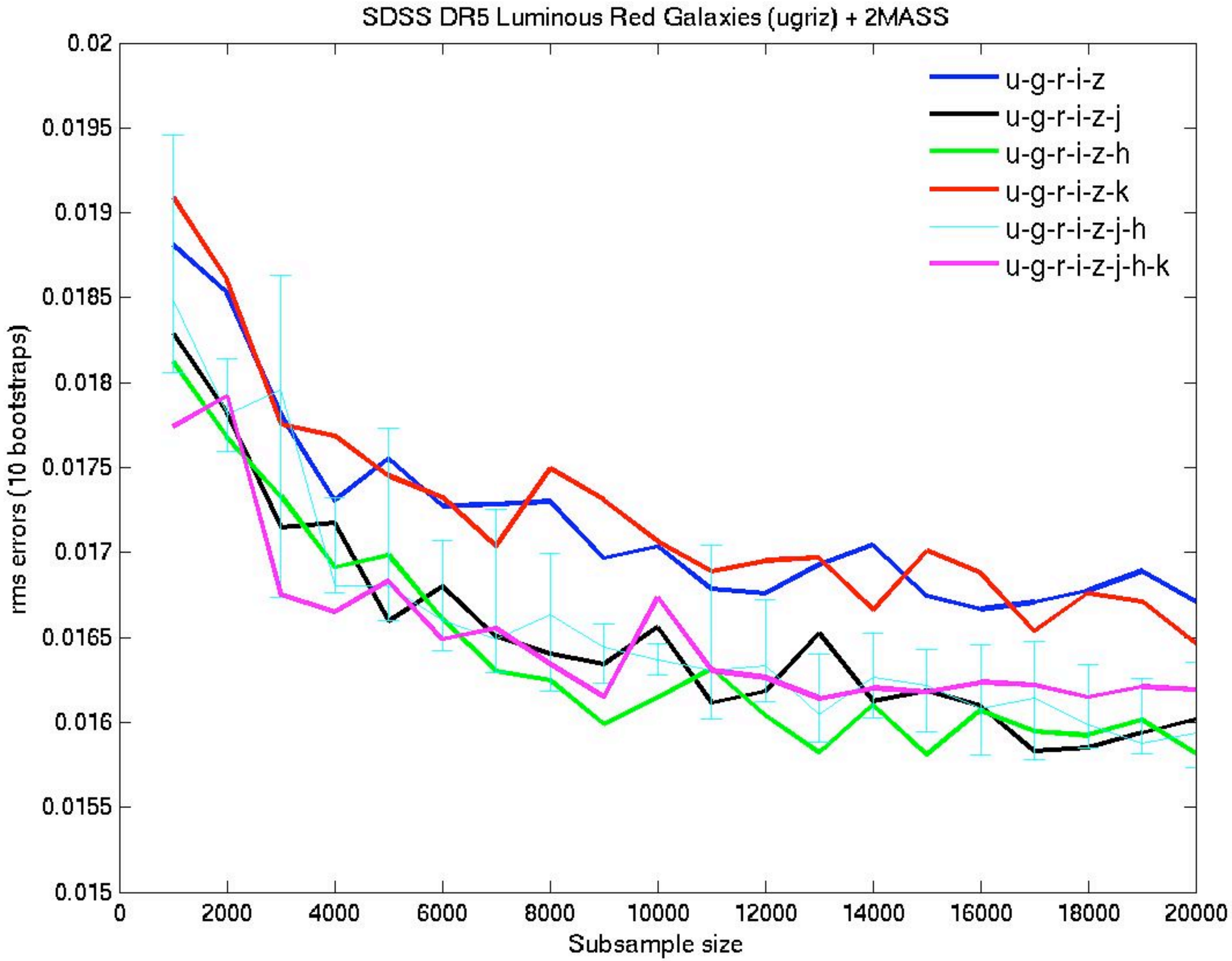
MGS + 2MASS 20000



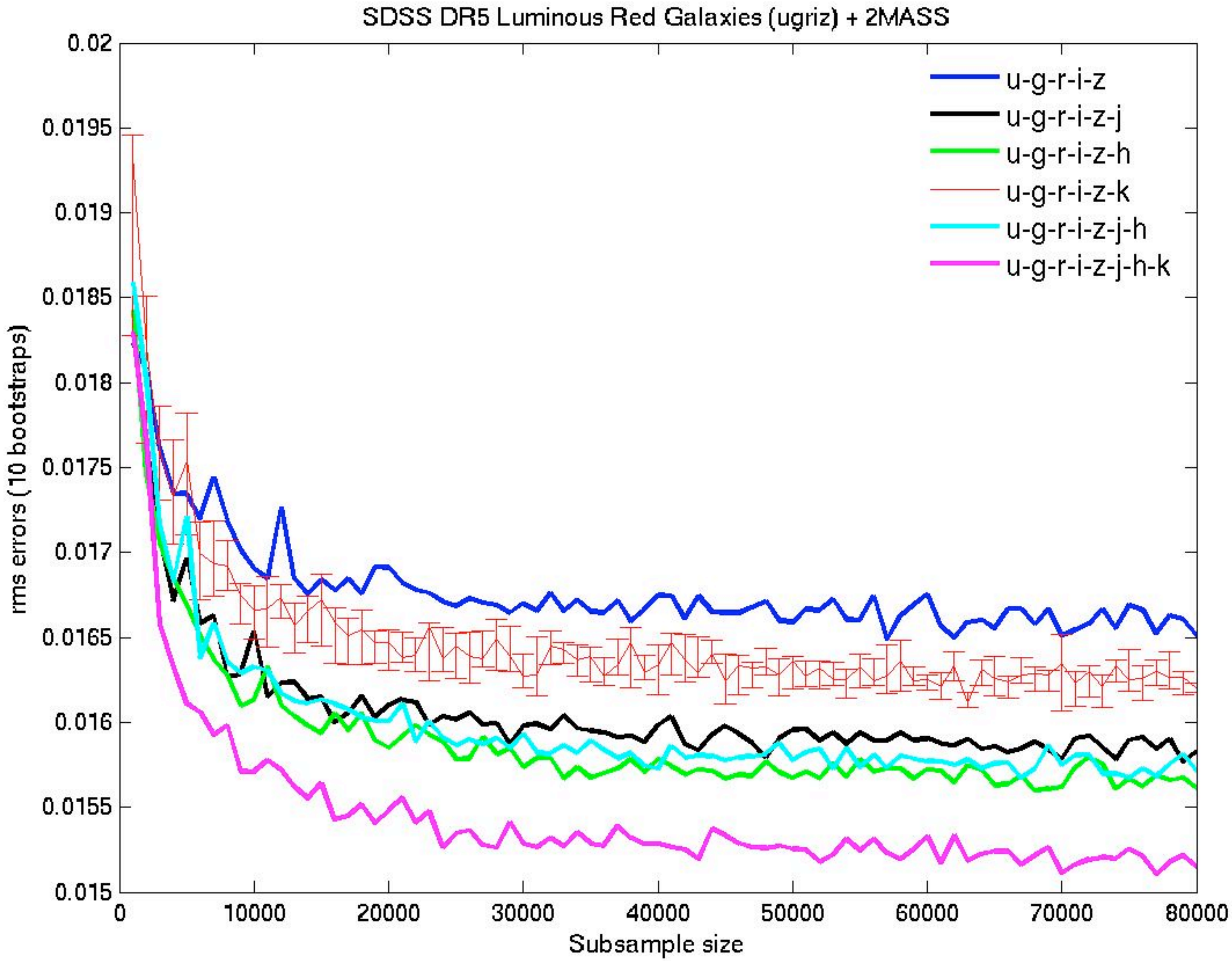
MGS + 2MASS 80000



LRG+ 2MASS 20000



LRG+ 2MASS 80000



Results?

- Morphology does not *generally* increase accuracy of photo-z estimation *with GPR*
- Better quality photometry and/or removal of error outliers does not help
- Additional Near IR filters (2MASS) increase accuracy
- Galaxy classification helps: MGS vs LRG
- Optimal filters? u-g-r-i-z and j-h-k