Laboratoire d'Astrophysique
Ecole Polytechnique Fédérale de Lausanne
Switzerland

http://lastro.epfl.ch

Kriging Interpolation

Application to the GREAT10 Star Challenge

Marc Gentile, Guldariya Nurbaeva,
Frederic Courbin, Georges Meylan

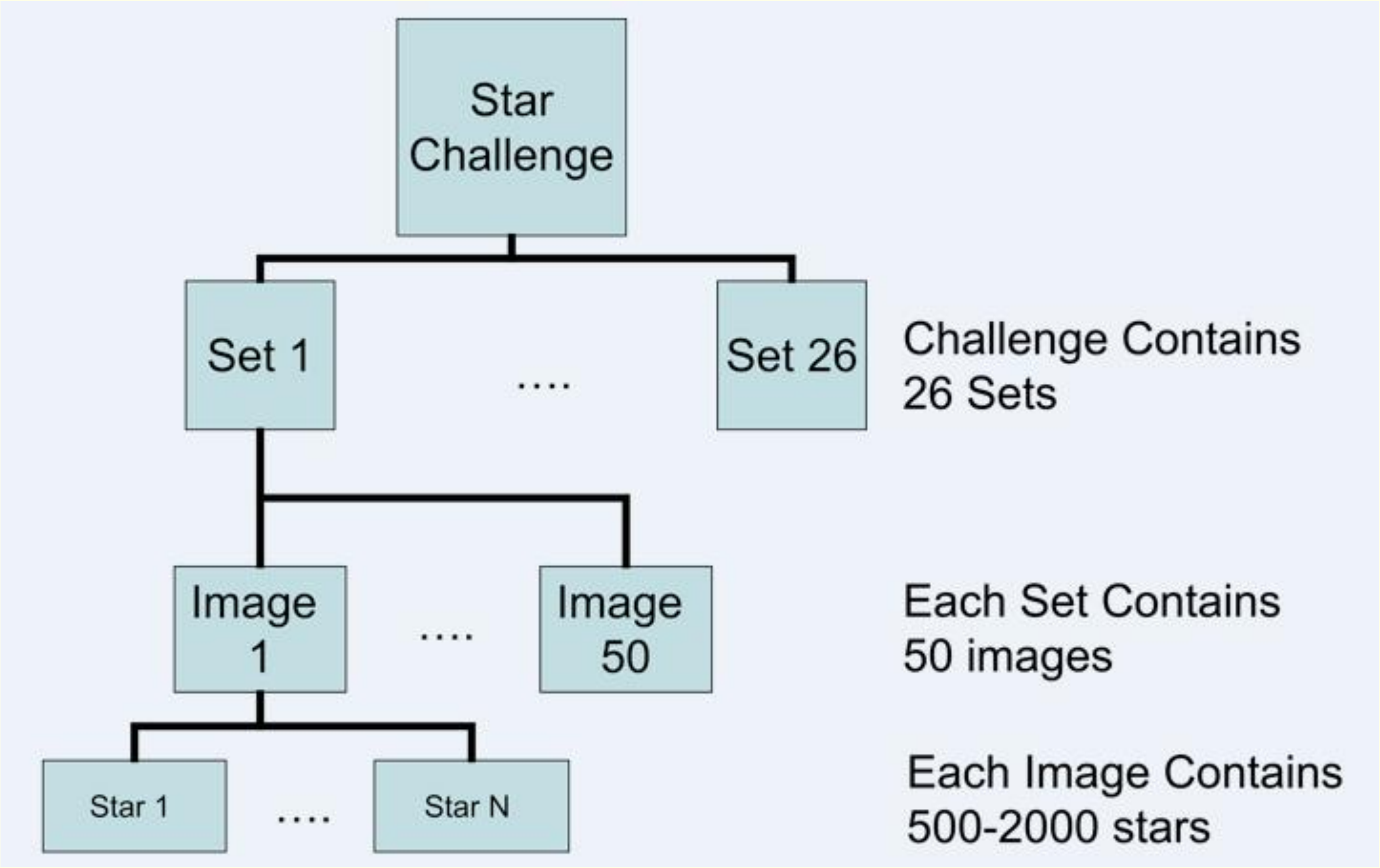
(Laboratoire d'Astrophysique de l'EPFL, Switzerland)

Agenda

- The GREAT10 Star Challenge
- Approaching the problem
- Introducing Kriging
- Some results on simulated Star Challenge data
- Some results on actual Star Challenge data

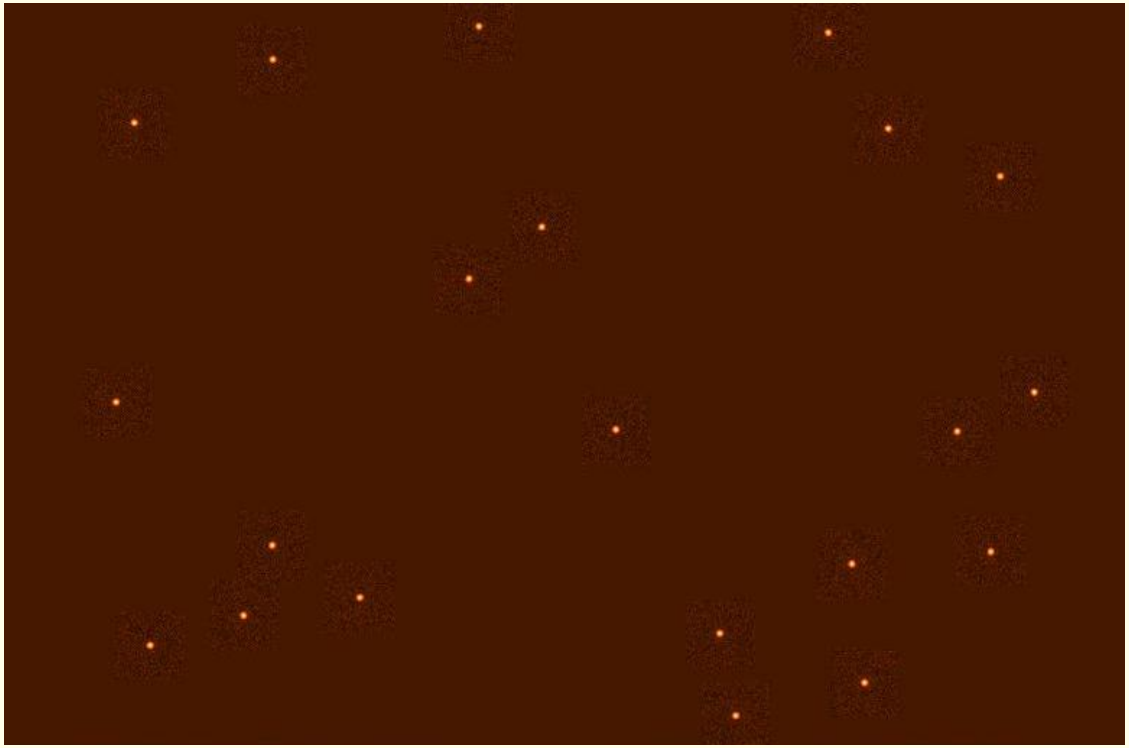
The GREAT10 Star Challenge

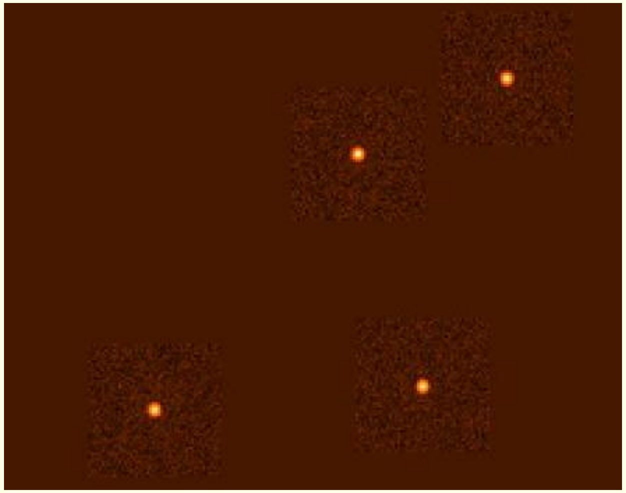
The data



The GREAT10 Star Challenge

The data



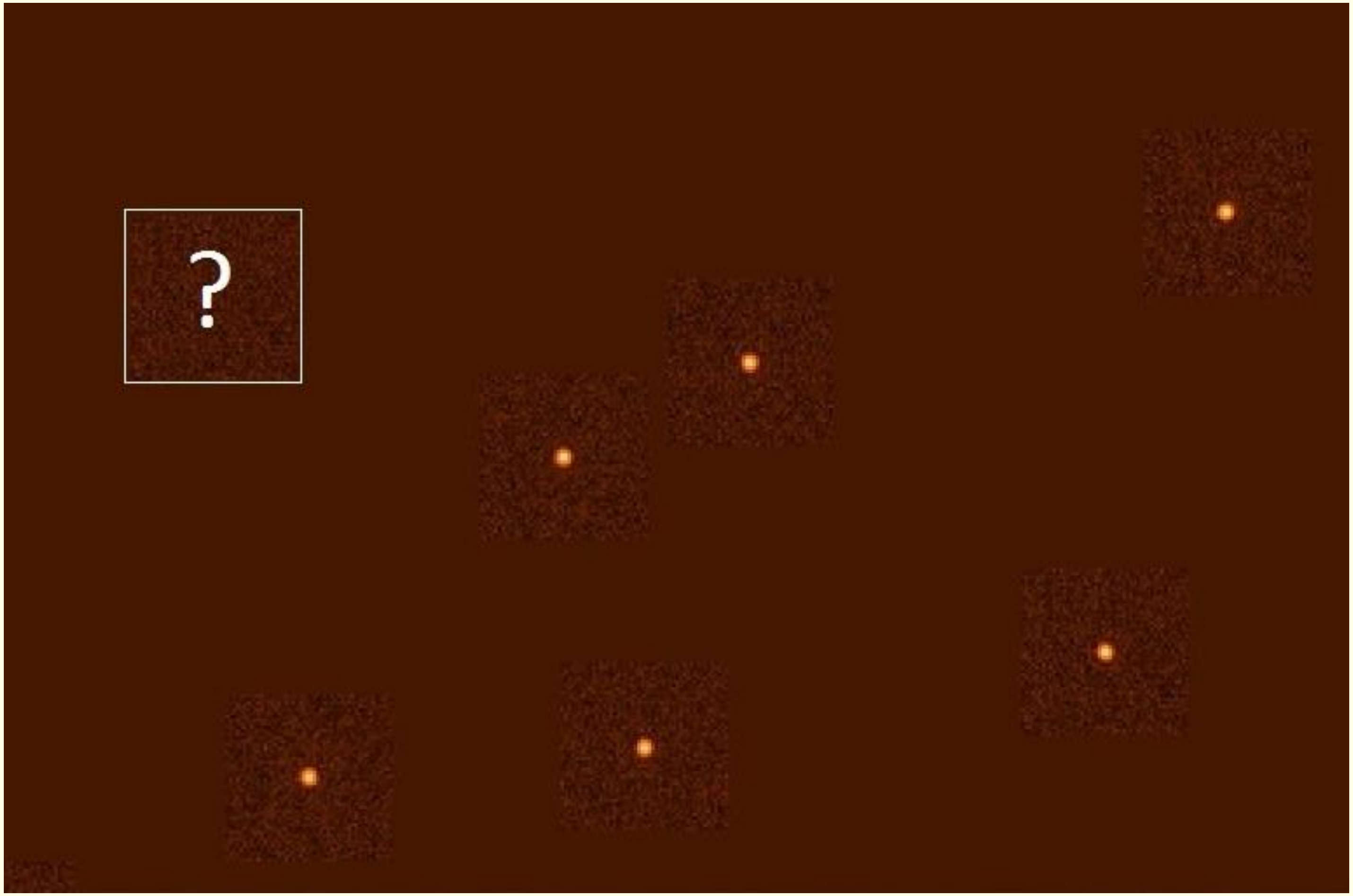


48x48 pixels square postage stamps

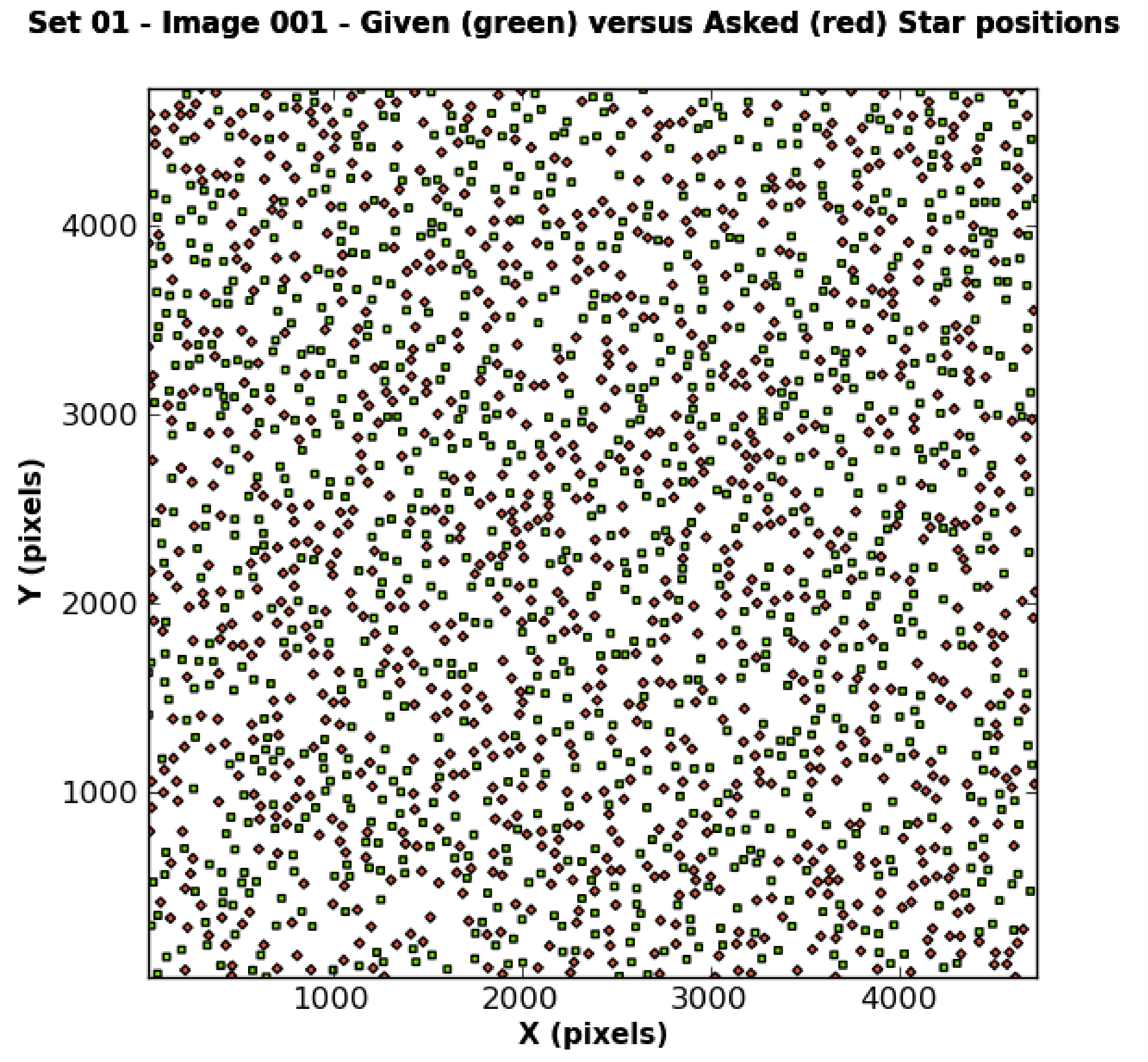
- Spatially-varying PSF field
- Images contain Gaussian noise

The GREAT10 Star Challenge

The challenge



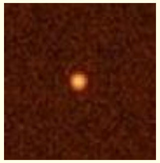
The GREAT10 Star Challenge



Approaching the problem

Key problems to tackle

How to predict the right star images at arbitrary non-star positions?



How to model individual star images?

? Spatial Prediction: Kriging

“Kriging“

- Over the past several decades, Kriging has become a fundamental spatial prediction tool in Geostatistics
  - geostatistics: a branch of statistics focusing on spatial or spatio-temporal data

- Found applications in many fields: mining, environmental sciences, hydrogeology, remote sensing… Why not astrophysics?

Spatial Prediction: Kriging

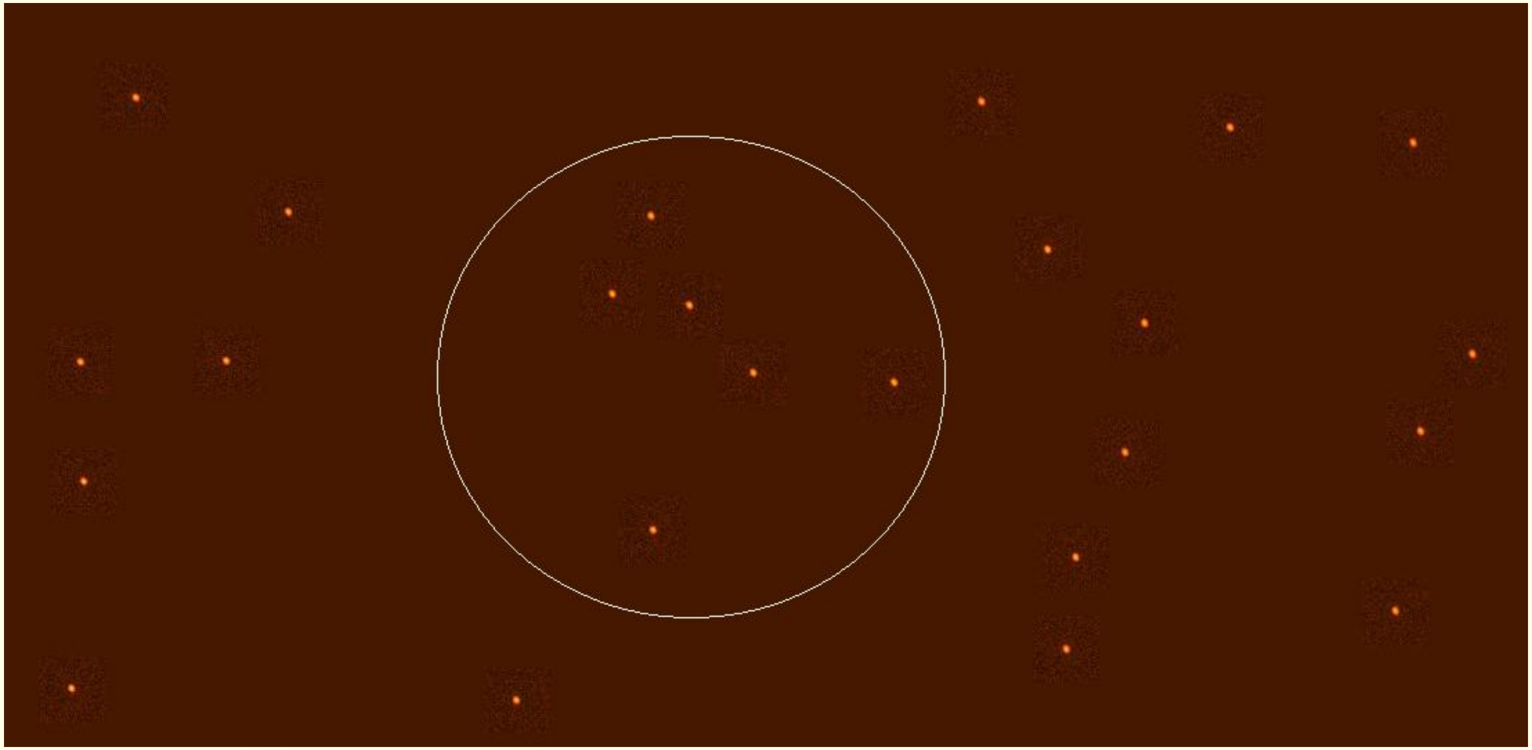
"Kriging"

- Named after mining engineer D. G. Krige, further developed by G. Matheron in the 60s
- A method for interpolating the value of a random field at an unobserved location based on available surrounding measurements
- The Kriging interpolation is local, exact and probabilistic

Spatial Prediction: Kriging

Kriging is a local interpolation method

- For its estimation, Kriging only considers measurements within some neighbourhood



Spatial Prediction: Kriging

Kriging is an exact interpolation method

- Predictions at known observed values yield exactly the same values again

- The interpolation produces a surface that passes exactly through all known points in the estimation neighbourhood

Spatial Prediction: Kriging

Kriging: a probabilistic interpolation method

- Assumes a random field of the form $Z(\mathbf{x}) = \mu(\mathbf{x}) + \varepsilon(\mathbf{x})$ with $E[\varepsilon(\mathbf{x})] = 0$ and covariance function $C(\mathbf{x}_j, \mathbf{x}_i) = E[\varepsilon(\mathbf{x}_i)\,\varepsilon(\mathbf{x}_j)]$

- Assumes the random field has some degree of stationary

- The prediction error can be estimated (Kriging variance)

Spatial Prediction: Kriging

Kriging assumes Intrinsic Stationarity

- Intrinsic stationarity
  - Constancy of the first two moments of the *differences*: $\boxed{[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})]}$

    $\boxed{E[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})] = 0}$ for small $\mathbf{h}$ at least

    $$\boxed{Var[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})] = E\left[(Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x}))^2\right] = 2\gamma(\mathbf{h})}$$

  - $\boxed{\gamma(\mathbf{h}) = \frac{1}{2}E\left[(Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x}))^2\right]}$ is the 'semi-variance'

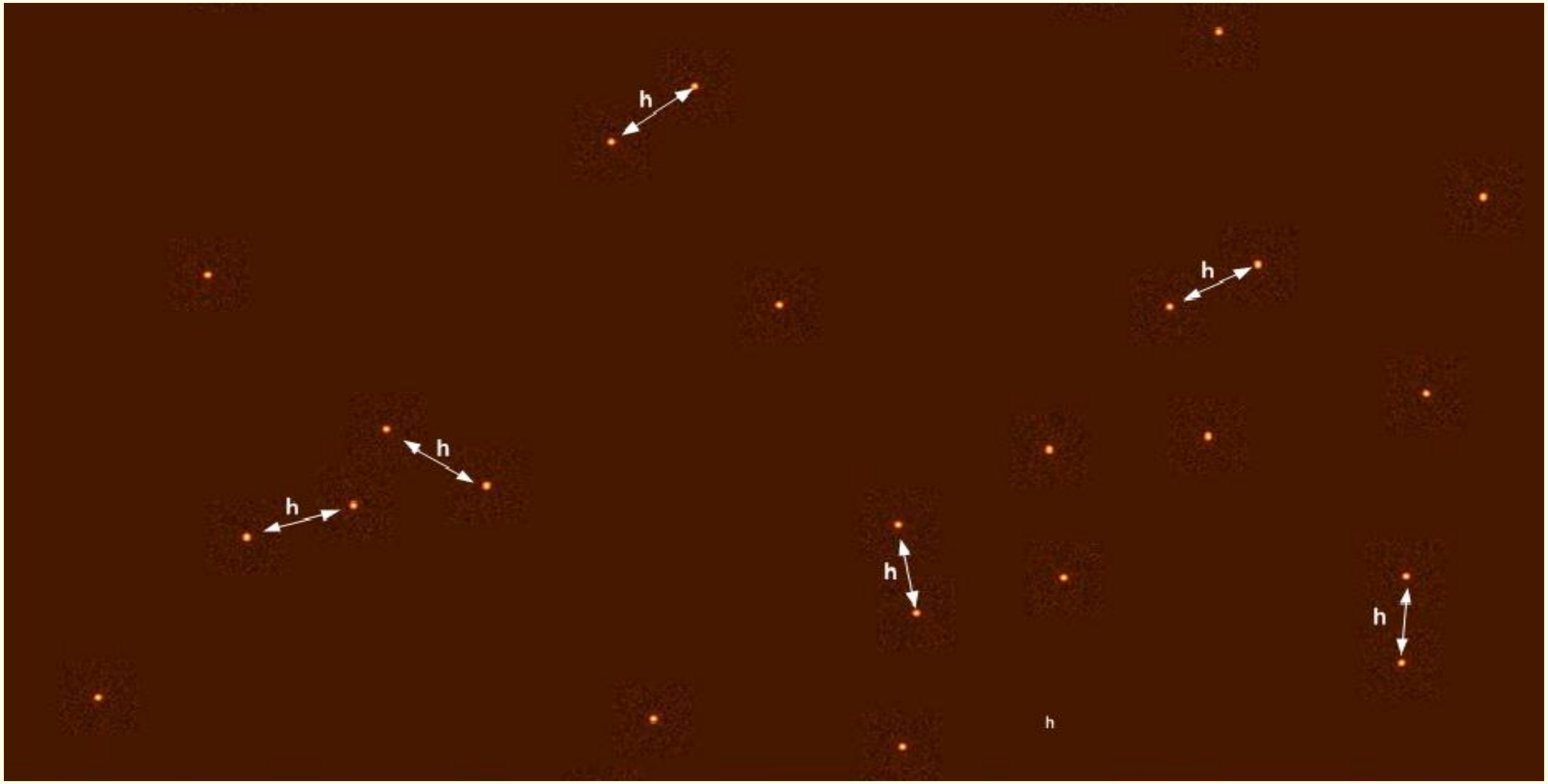
  - $\boxed{h = \mathbf{x}_j - \mathbf{x}_i}$ is the 'lag' separation distance

Spatial Prediction: Kriging

Semi-Variogram

- Compute the semi-variance for a range of lags

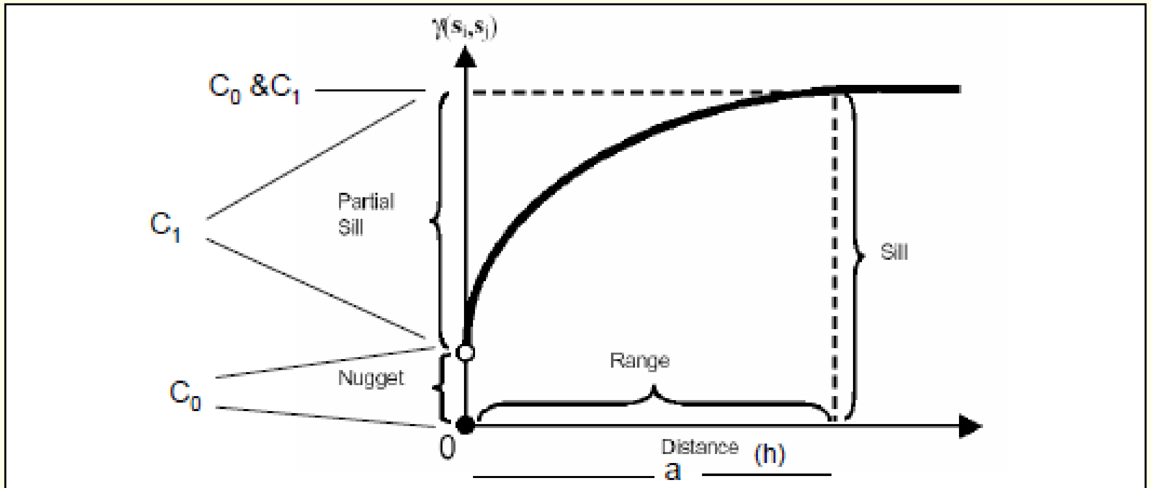
$$\boxed{h = \mathbf{x}_j - \mathbf{x}_i}$$



Spatial Prediction: Kriging

Semi-Variogram

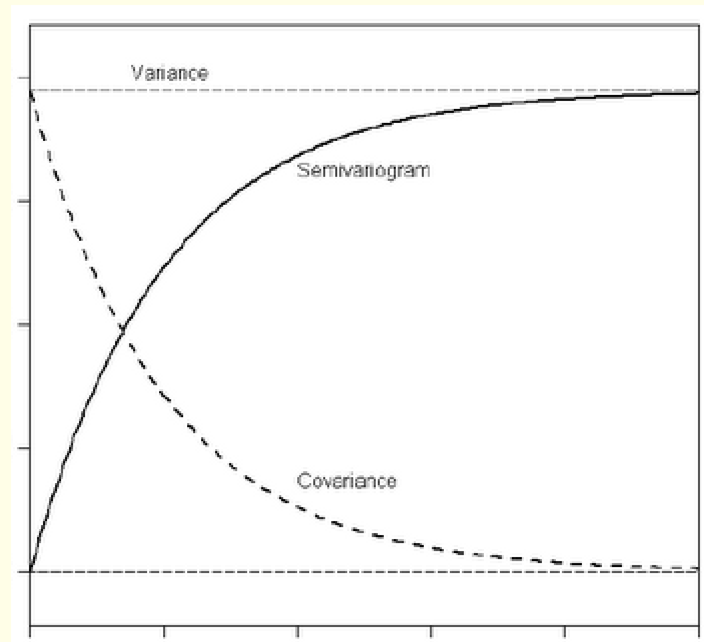
- Compute the semi-variance for a range of lags $\boxed{h = \mathbf{x}_j - \mathbf{x}_i}$ … one gets some like this:



Nugget (C_0): Semivariance at distance 0: Represents microscale variations or measurement errors

Range (a) : Distance at which semivariance levels off: Represents the spatially correlated portion

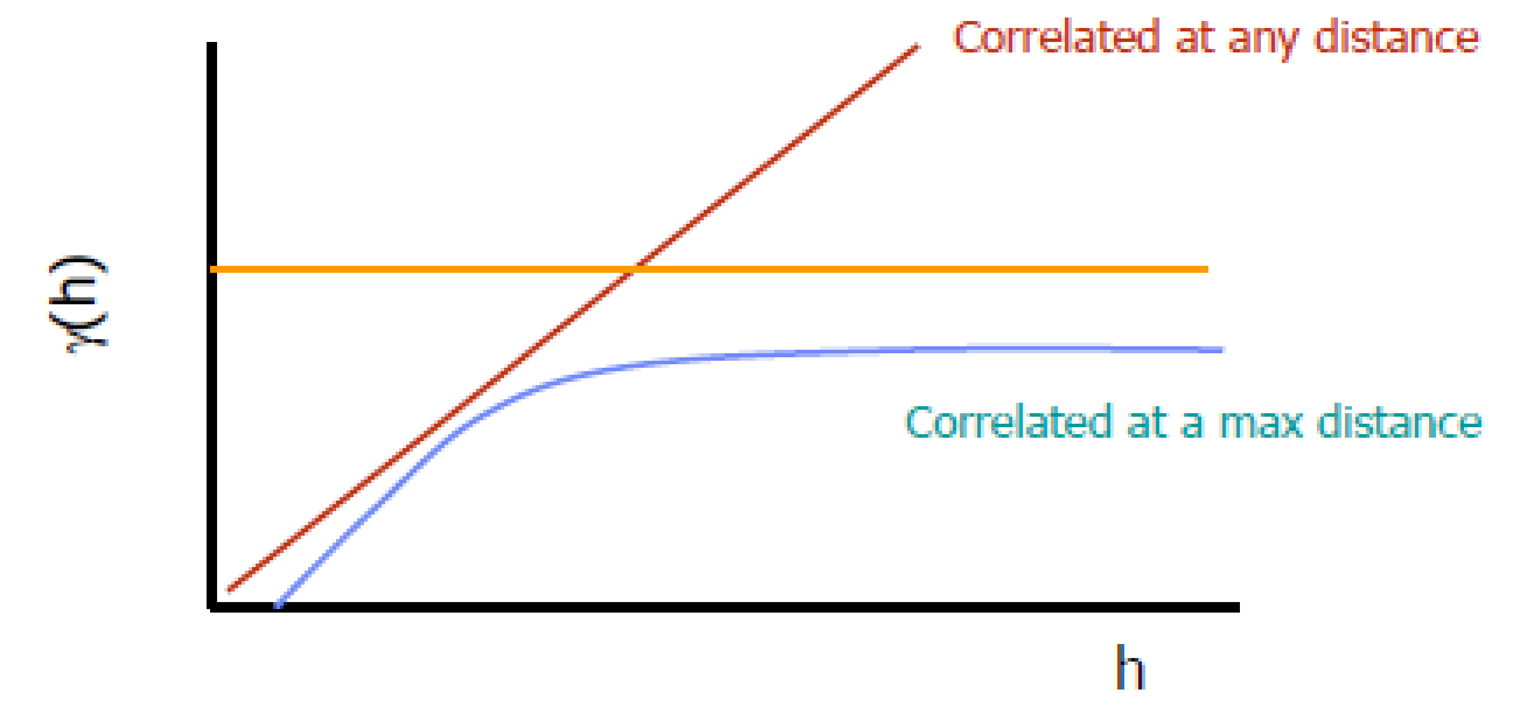
Sill ($C_0 + C_1$): Semivariance at which leveling takes place



Spatial Prediction: Kriging

Semi-Variogram

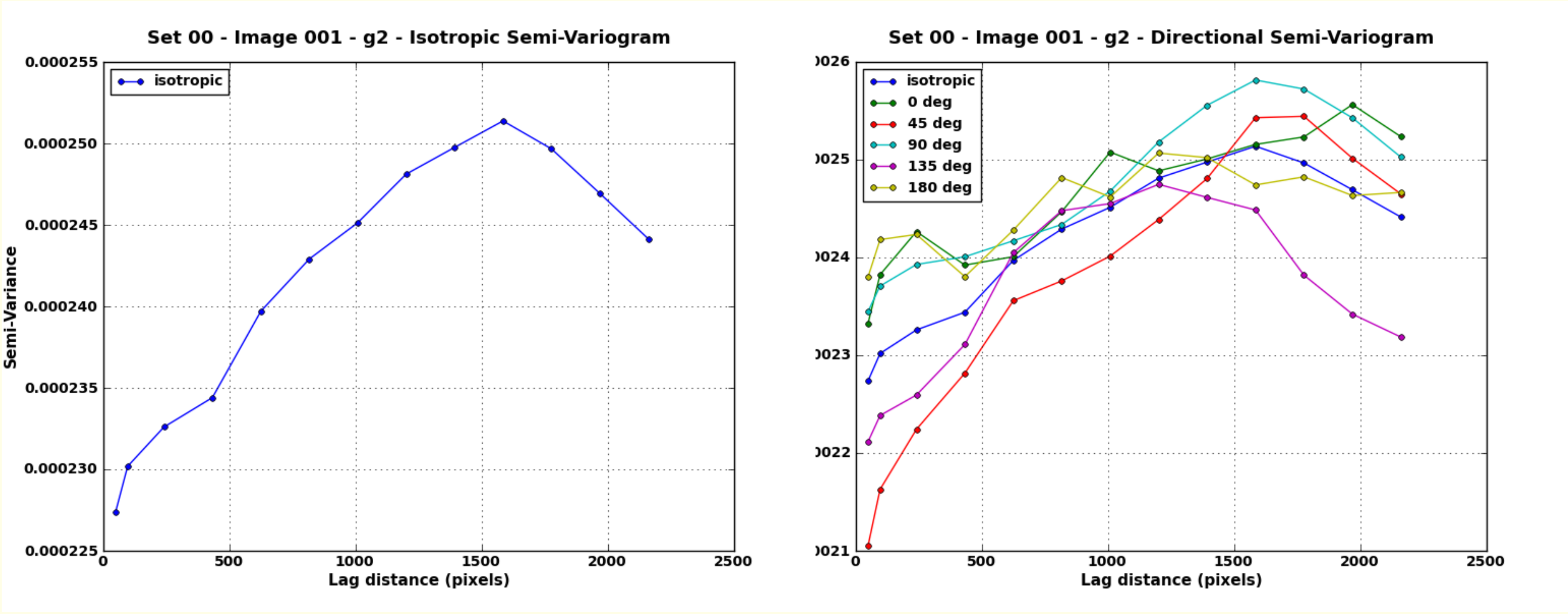
- Estimate the correlation distance in the data



Spatial Prediction: Kriging

Sample GREAT10 Galaxy Challenge:

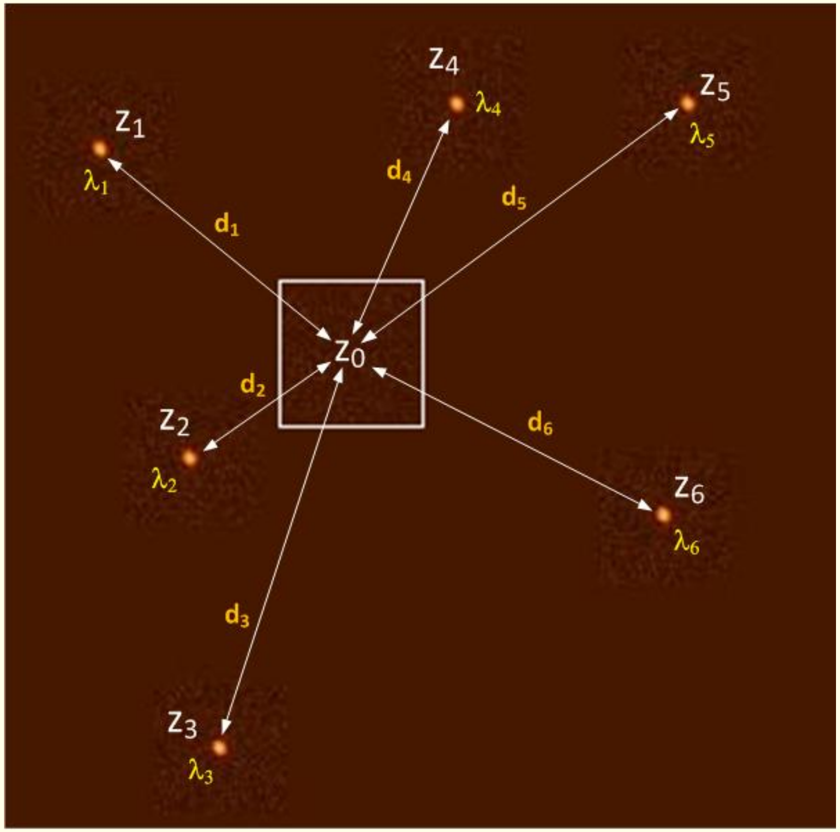
- Training Set Shear Semi-Variogram



Spatial Prediction: Kriging

Kriging prediction

- Made using a weighted linear estimator



$$z_0^* = z^*(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i \, z(\mathbf{x}_i)$$

λ_i: weight at location $\mathbf{x}_i$

z_0^*: predicted value at $\mathbf{x}_0$

d_i: distance between $\mathbf{x}_i$ and $\mathbf{x}_0$

N: number of sample values used in prediction

Spatial Prediction: Kriging

Kriging prediction: computing the weights

- Kriging aims at minimizing the so-called Kriging Variance

$$Var\left(Z_0^*\right) = E\left[\left(Z^*\left(\mathbf{x}_0\right) - Z\left(\mathbf{x}_0\right)\right)^2\right] = 2\sum_{i=1}^{N}\lambda_i\,\gamma\left(\mathbf{x}_i,\mathbf{x}_0\right) - \sum_{i=1}^{N}\lambda_i\lambda_j\gamma\left(\mathbf{x}_i,\mathbf{x}_j\right)$$

- Subject to the unbiaseness condition

$$\boxed{E\left[Z^*\left(\mathbf{x}_0\right) - Z\left(\mathbf{x}_0\right)\right] = \sum_{i=1}^{N}\lambda_i\,\mu\left(\mathbf{x}_i\right) - \mu\left(\mathbf{x}_0\right) = 0}$$

Spatial Prediction: Kriging

Ordinary Kriging Equations

- Ordinary Kriging (OK): most commonly-used
- Ordinary Kriging assumes $\boxed{Z(\mathbf{x}) = \mu(\mathbf{x}) + \varepsilon(\mathbf{x})}$

$\boxed{\mu(\mathbf{x}) = \mathbf{E}[Z(\mathbf{x})] = \mu \quad \textbf{constant but unknown}}$

which implies $\boxed{\sum_{i=1}^{N} \lambda_i = 1}$

- Ordinary Kriging Equations

$$\boxed{\sum_{j=1}^{N} \lambda_j \gamma(\mathbf{x}_j - \mathbf{x}_i) + m = \gamma(\mathbf{x}_0 - \mathbf{x}_j) \qquad i = 1, .., N} \qquad \boxed{\sum_{i=1}^{N} \lambda_i = 1}$$

m: **Lagrange multiplier**

Spatial Prediction: Kriging

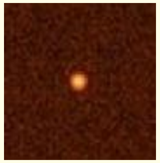
Main Kriging Steps

- Compute the semi-variance → semi-variogram
- Fit experimental variogram → variogram model
- At unknown location estimate Kriging weights within some neighbourhood
- Solve the Kriging equation to find the weights λ_i
- Kriging estimate is: $\boxed{z_0^* = z^*(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i \, z(\mathbf{x}_i)}$

Approaching the problem

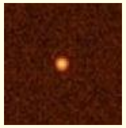
Key problems to tackle

How to predict the right star images at arbitrary non-star positions?



How to model individual star images?

Approaching the problem

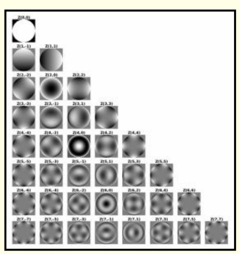


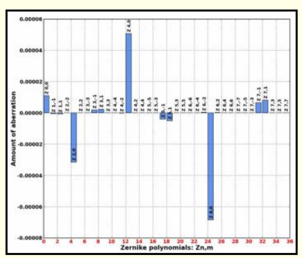
Modeling individual images

- Fit against a realistic star model (Moffat …)

$$p(r) = \frac{\beta - 1}{\pi\alpha^2}\left[1 + \left(\frac{r}{\alpha}\right)^2\right]^{-\beta}$$

- Mathematical decomposition in some basis (wavelets, Zernike polynomials …)



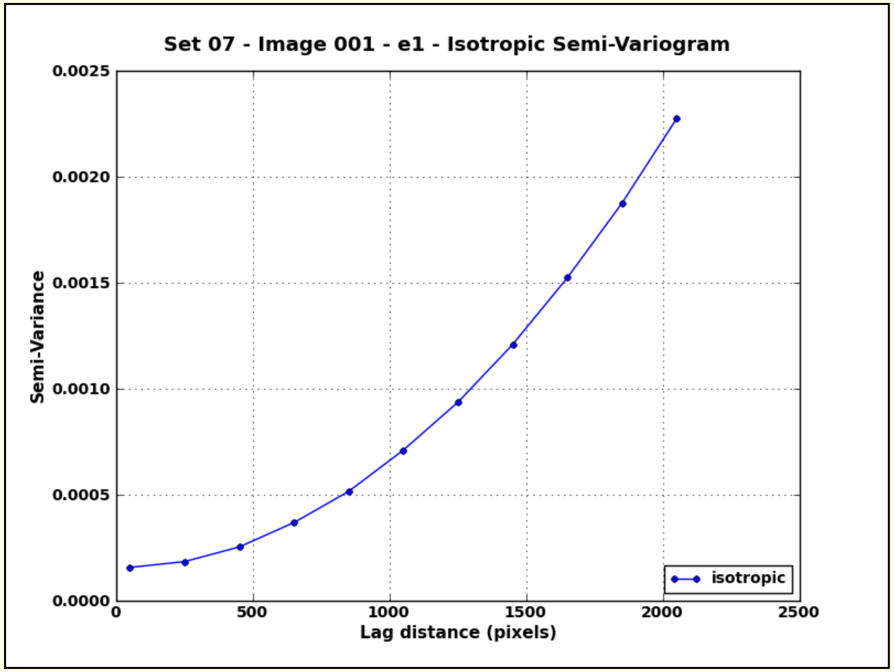


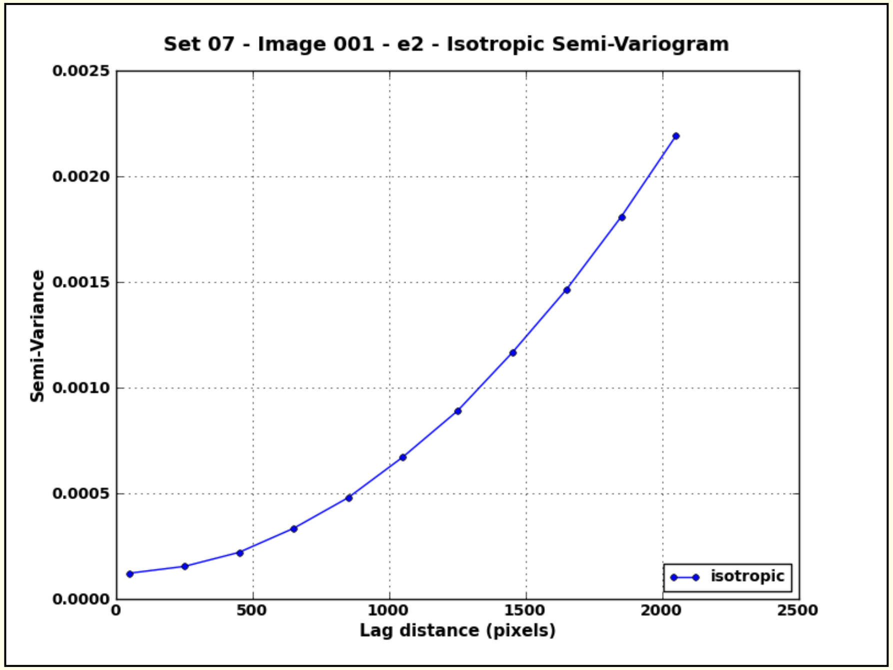
- Direct use of pixels: individually, by block …

Kriging applied to the Star Challenge

Sample Experimental Semi-Variograms

- Point Spread Function ellipticity (e_1,e_2)

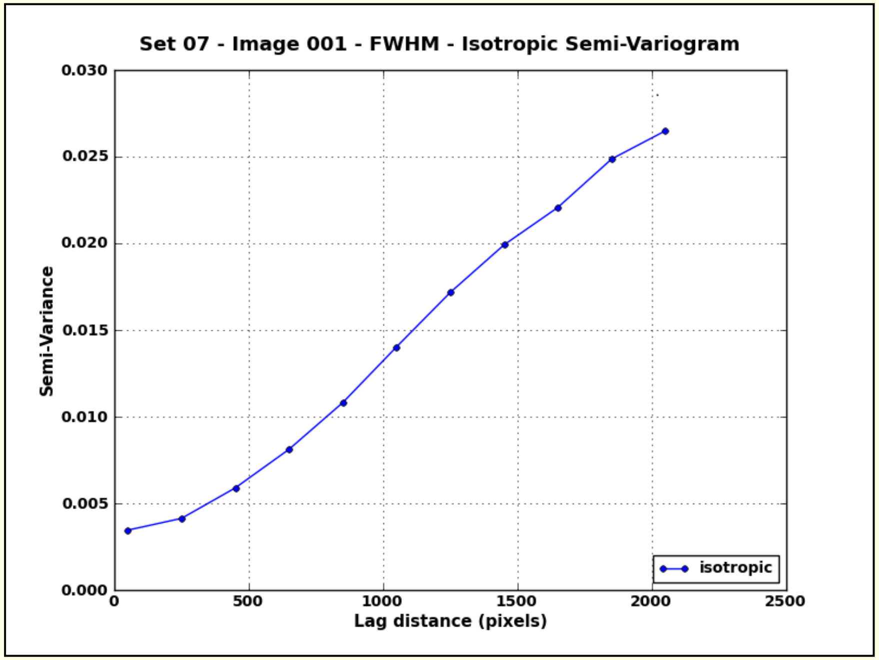


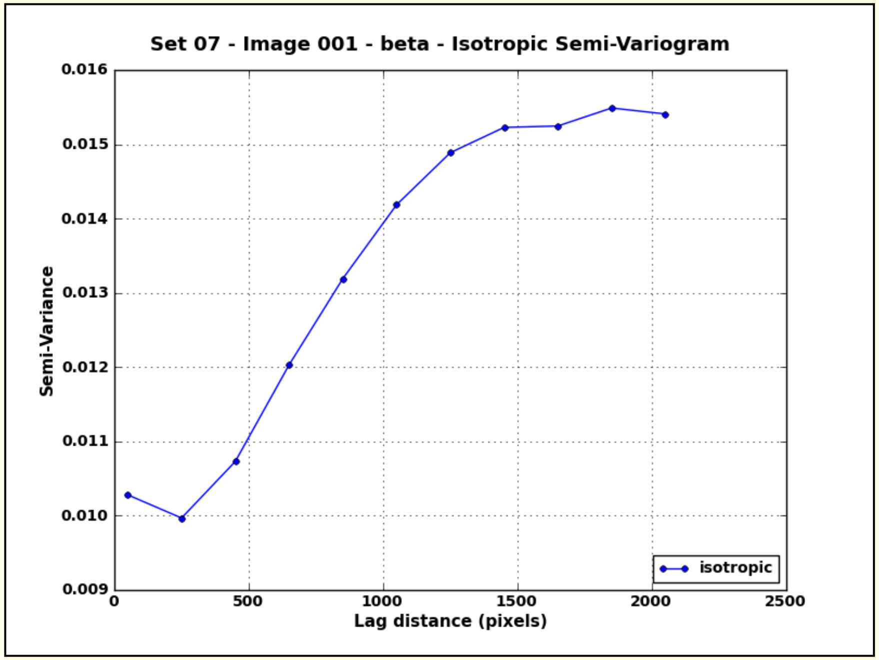


Kriging applied to the Star Challenge

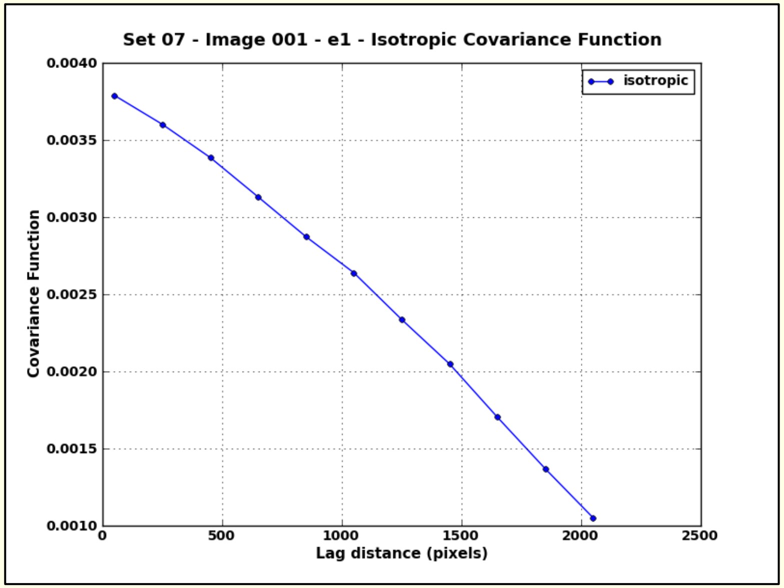
Sample Experimental Semi-Variograms

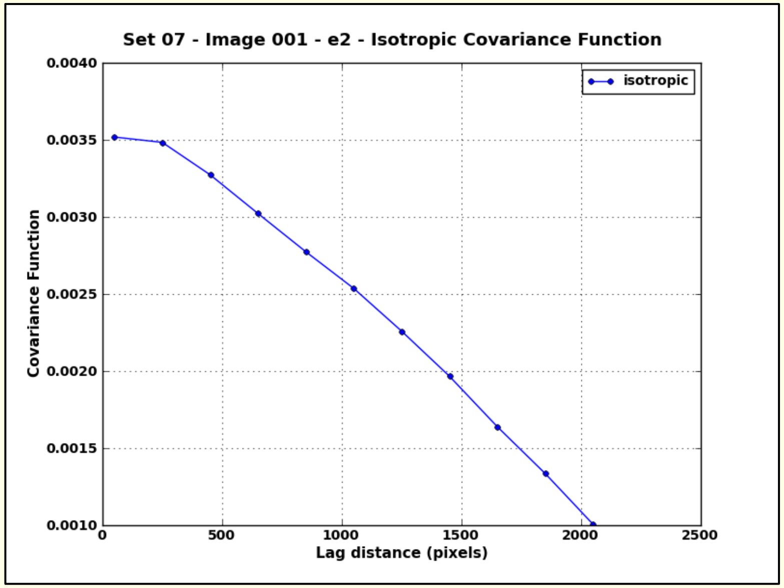
- Point Spread Function FWHM and Moffat beta

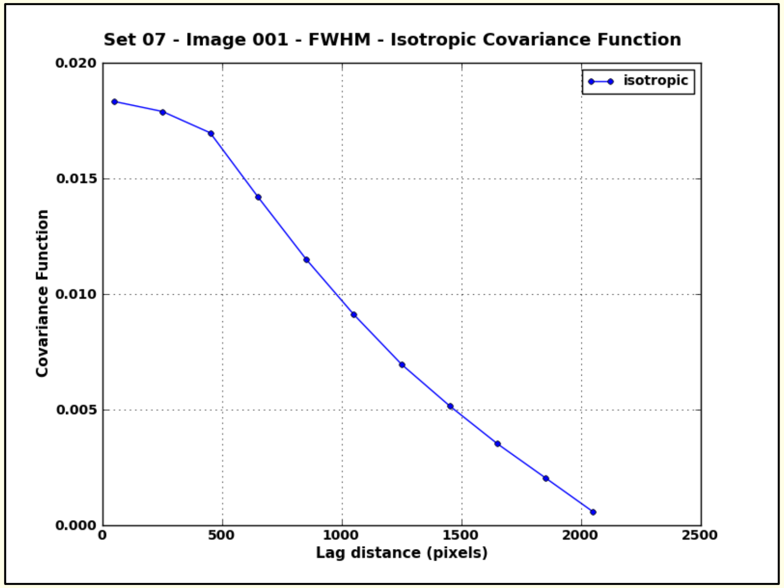


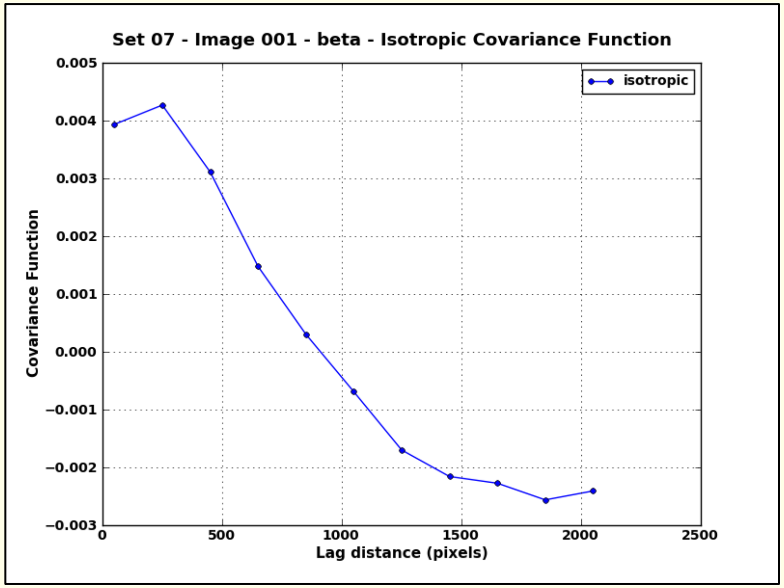


Sample Experimental Covariance Functions

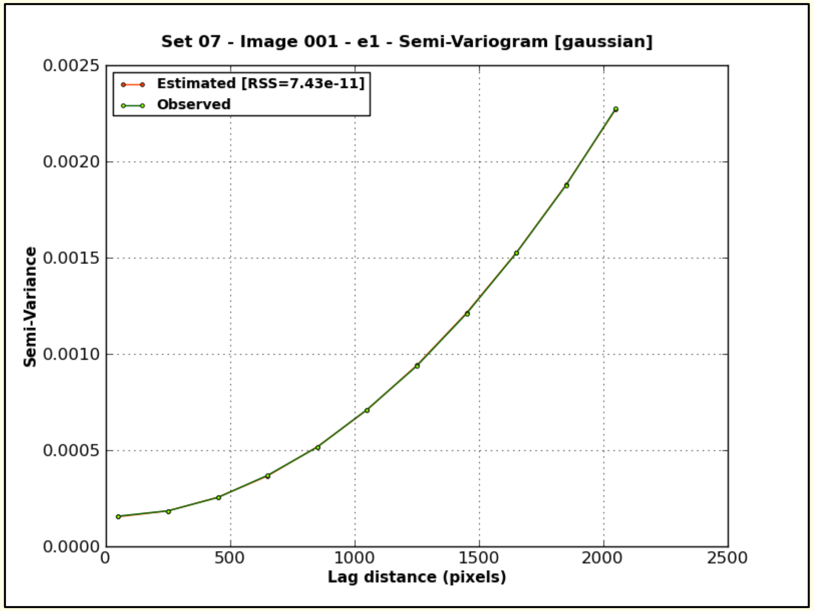


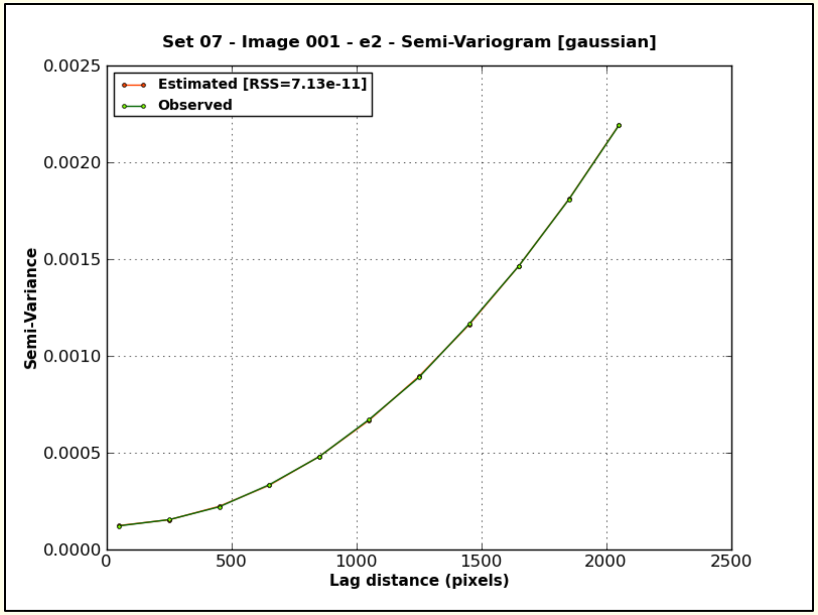


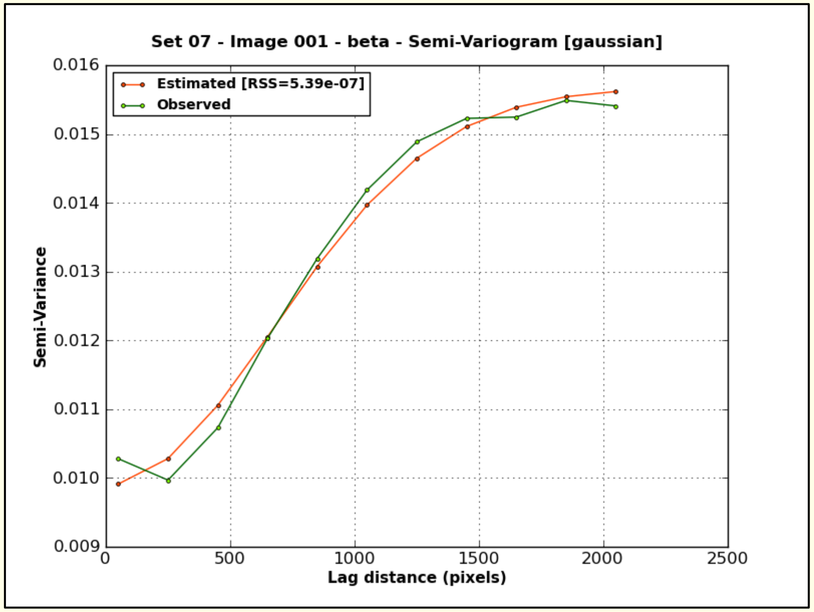


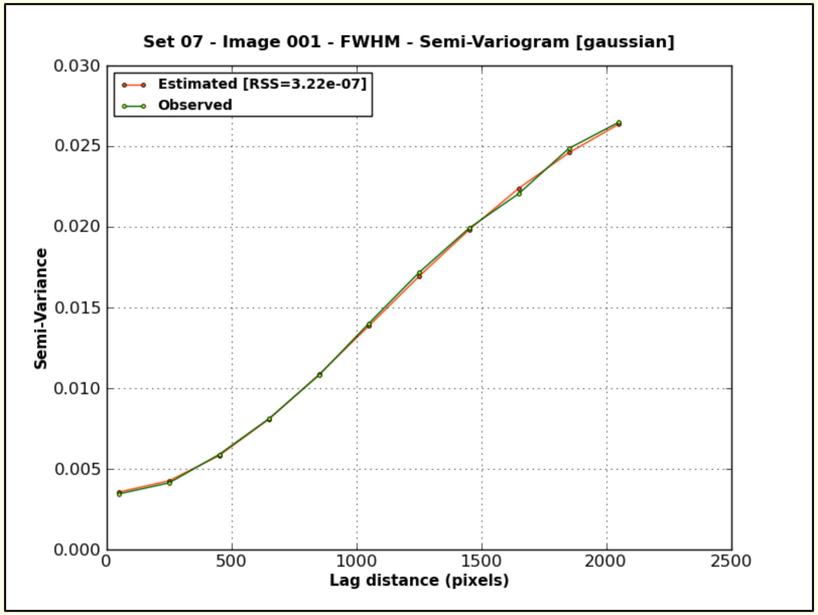


Sample Model-Fitted Semi-Variogram



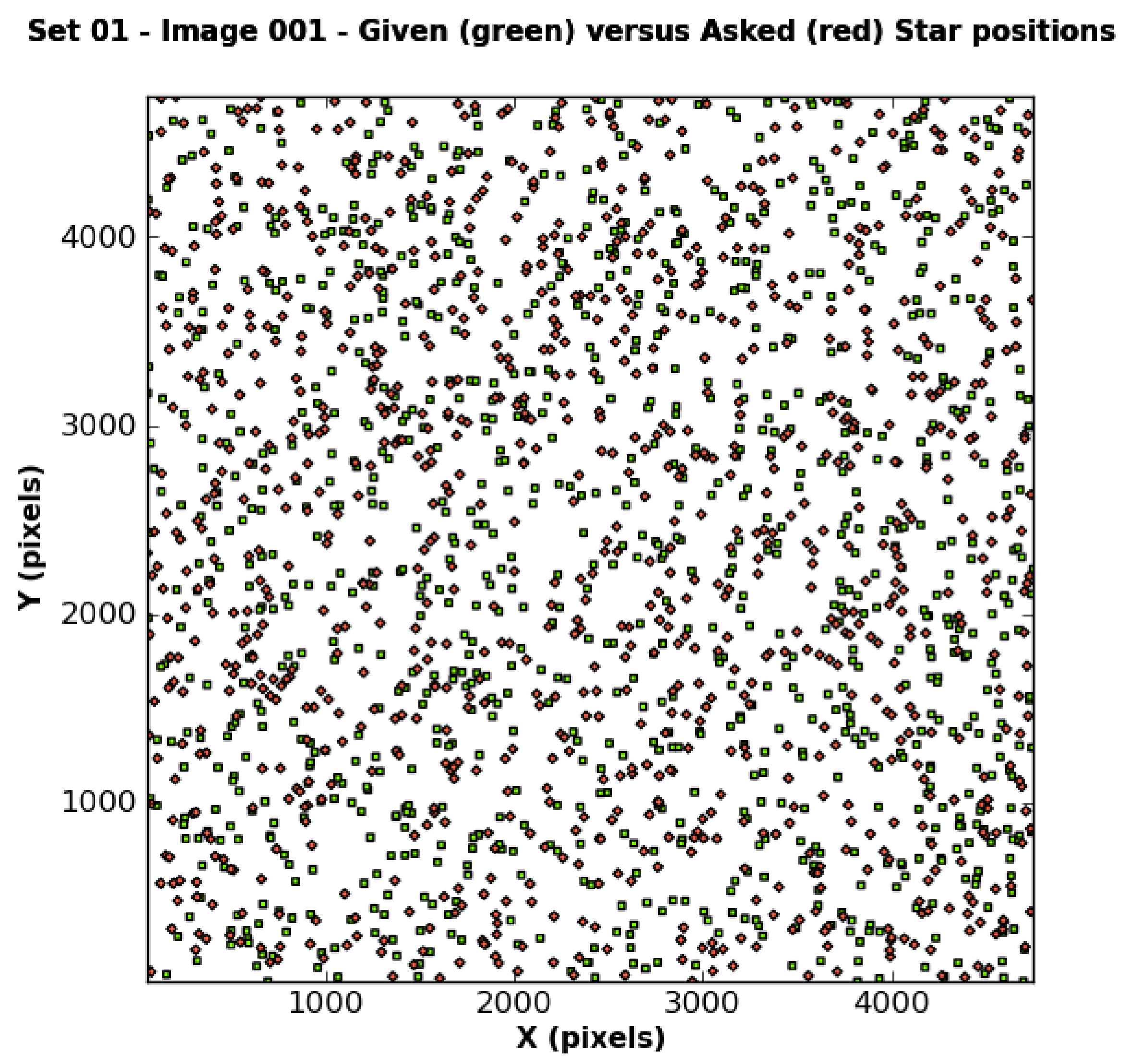




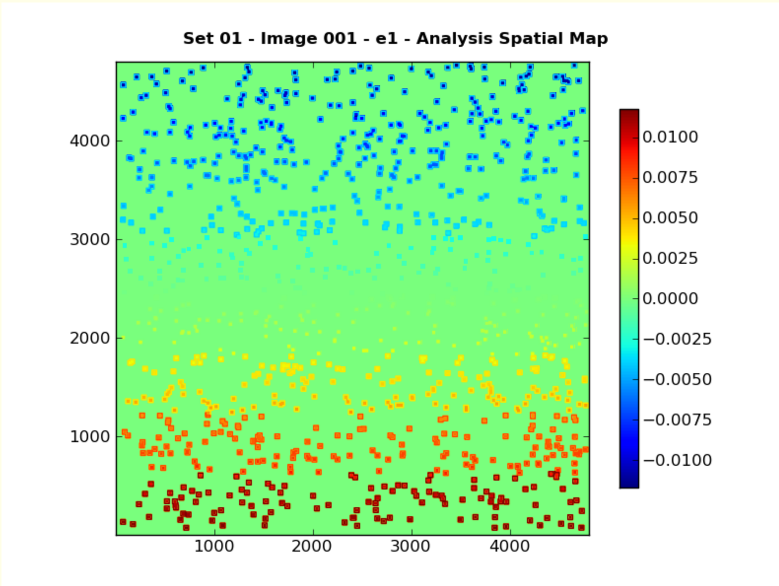


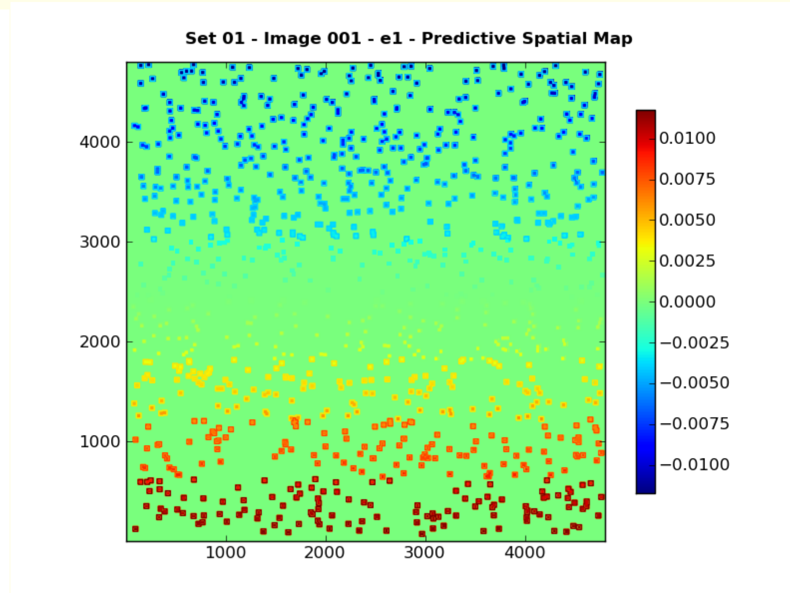
Kriging on Star Challenge-like Simulated Data

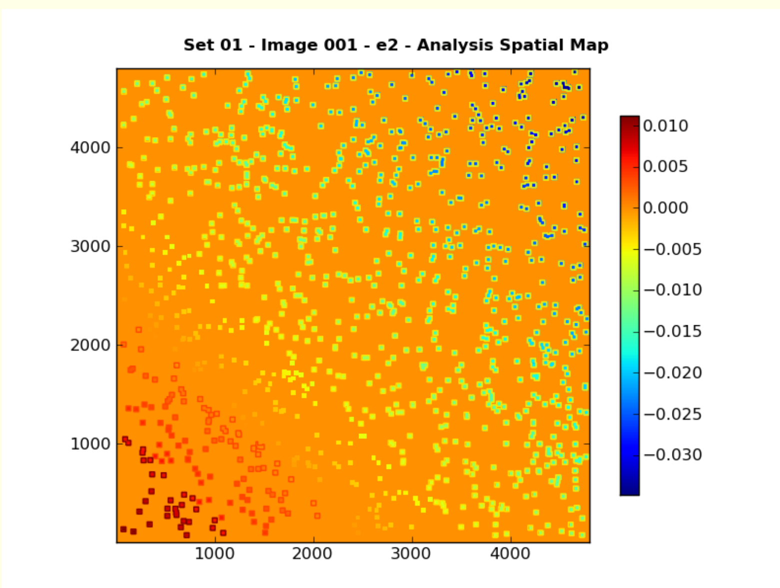
Ellipticity: predictions at asked positions

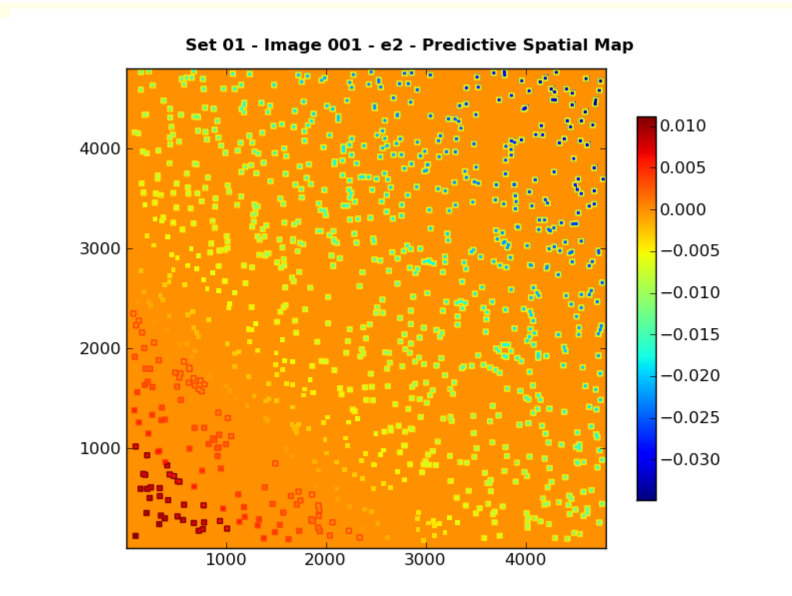


Kriging from Simulated Images: predictions

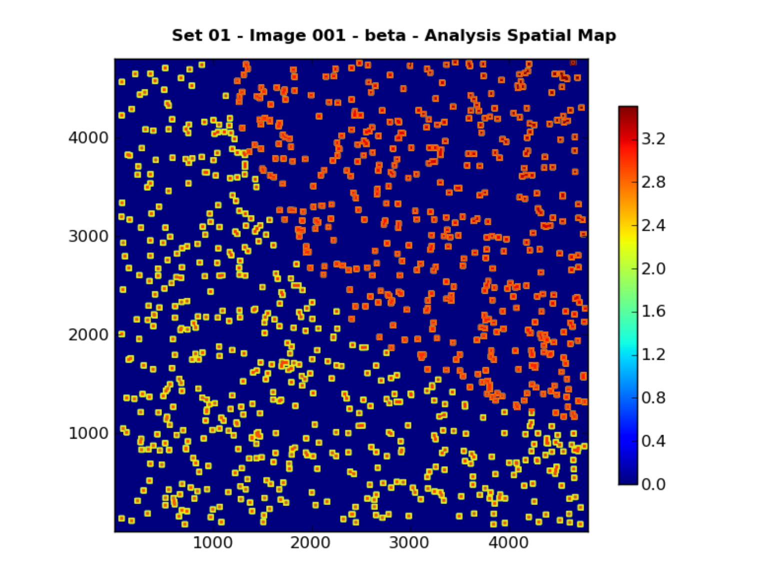


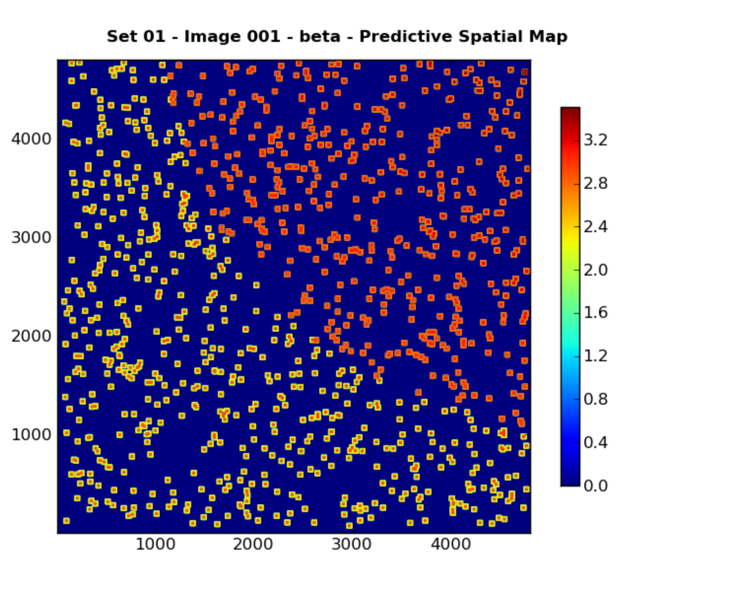


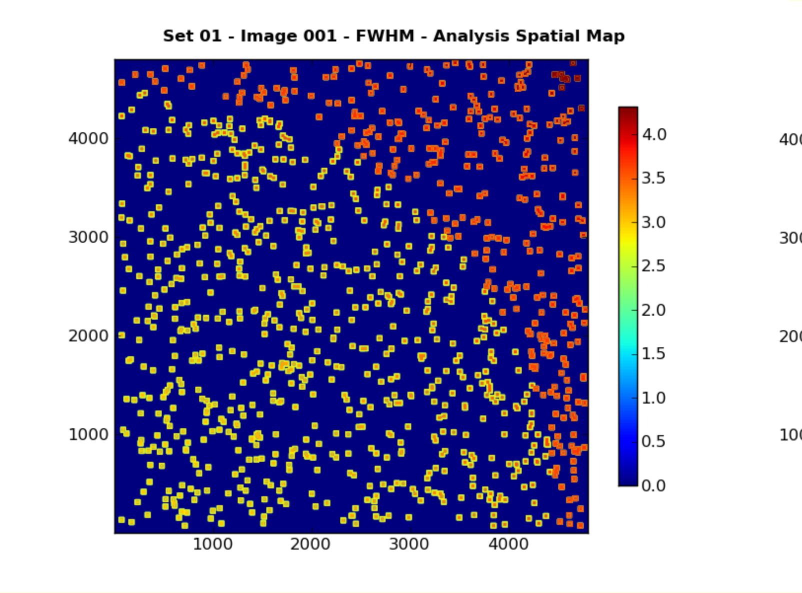


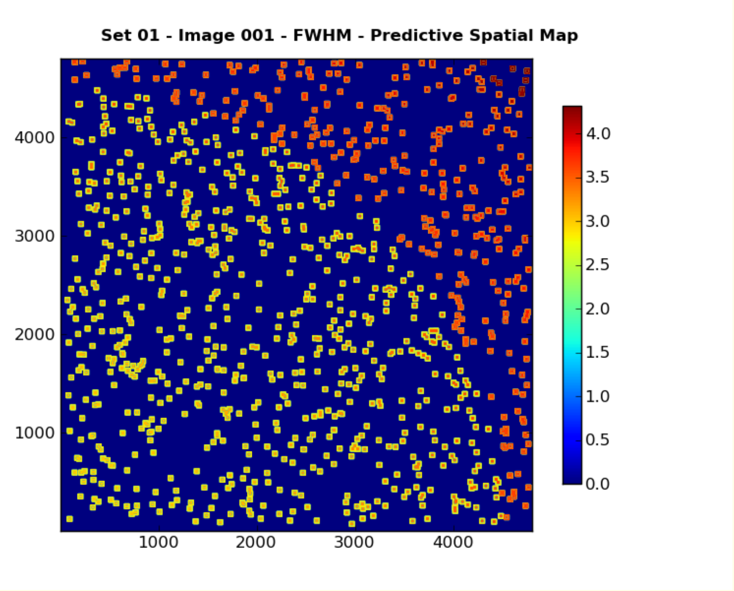


Kriging from Simulated Images: predictions

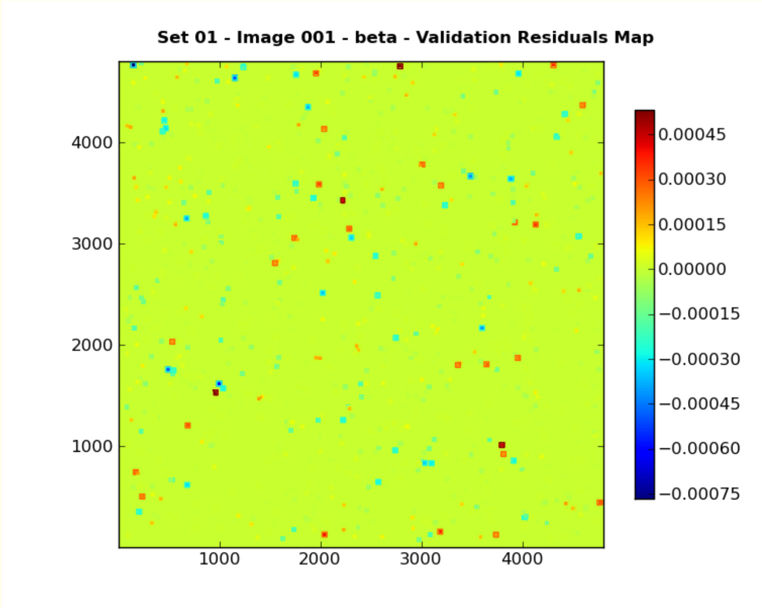


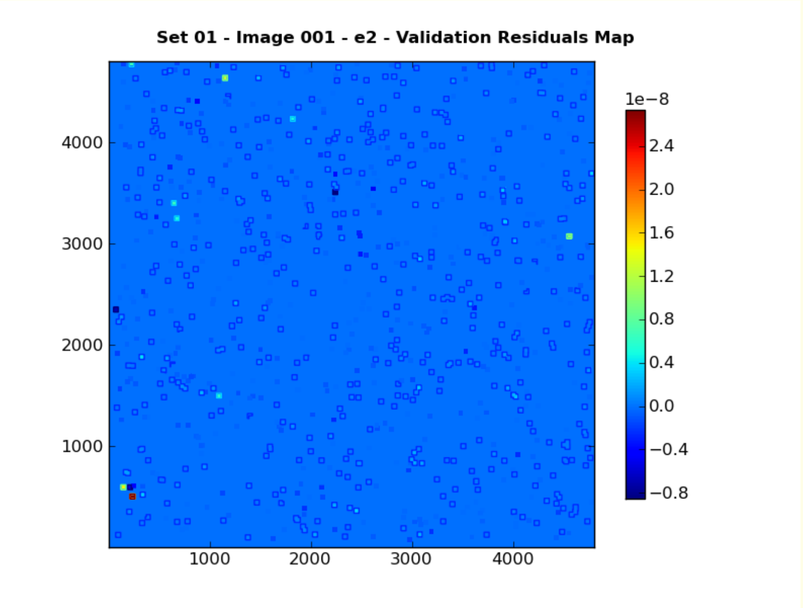


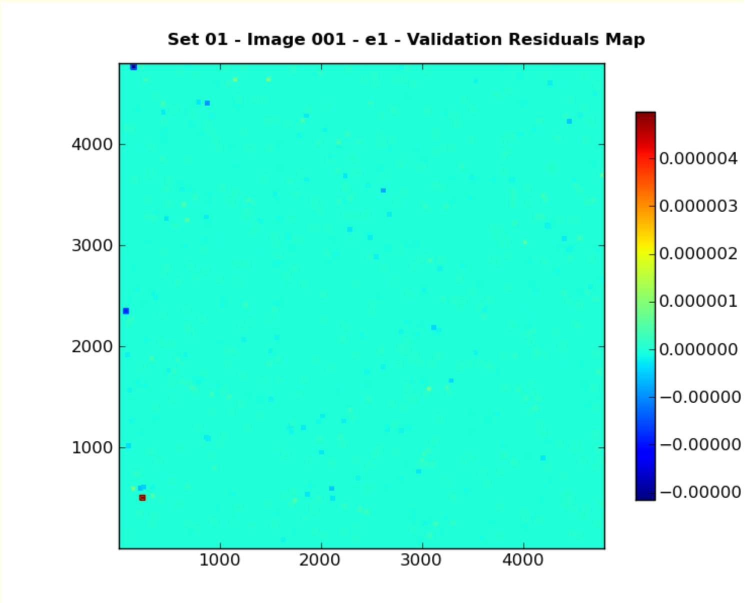


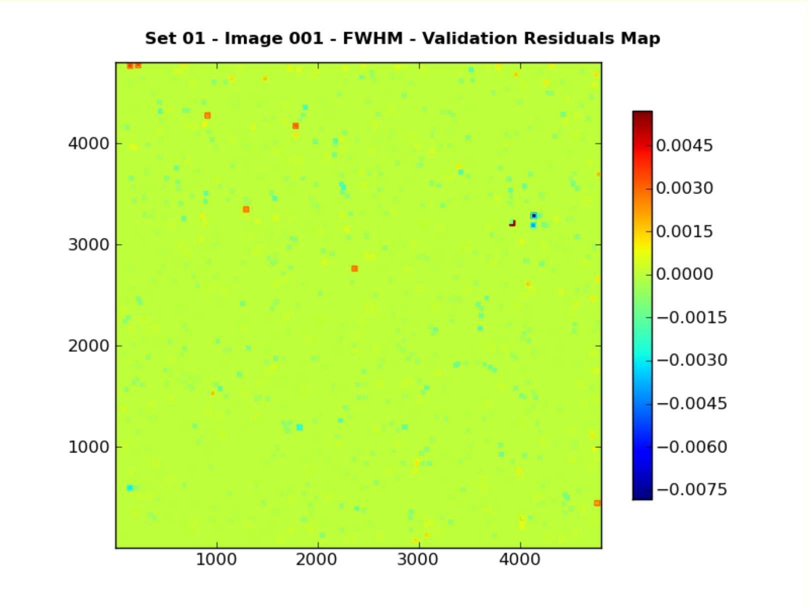


Kriging from Simulated Images: residuals



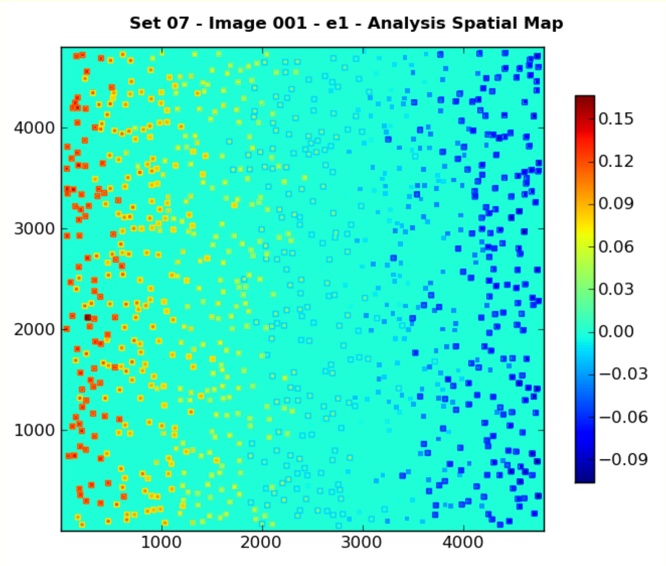


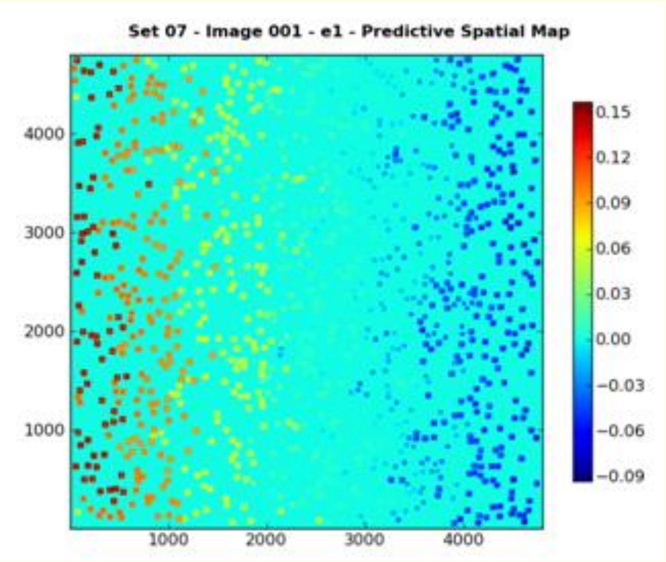


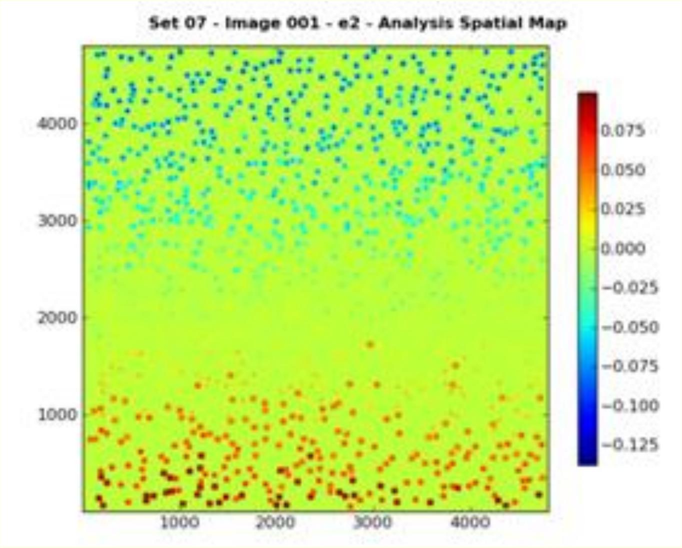


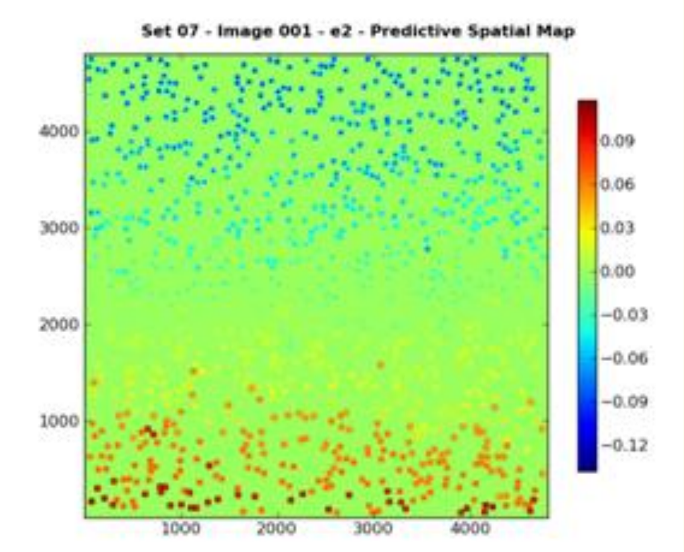
Kriging on actual Star Challenge Data

Ellipticity: predictions at asked positions

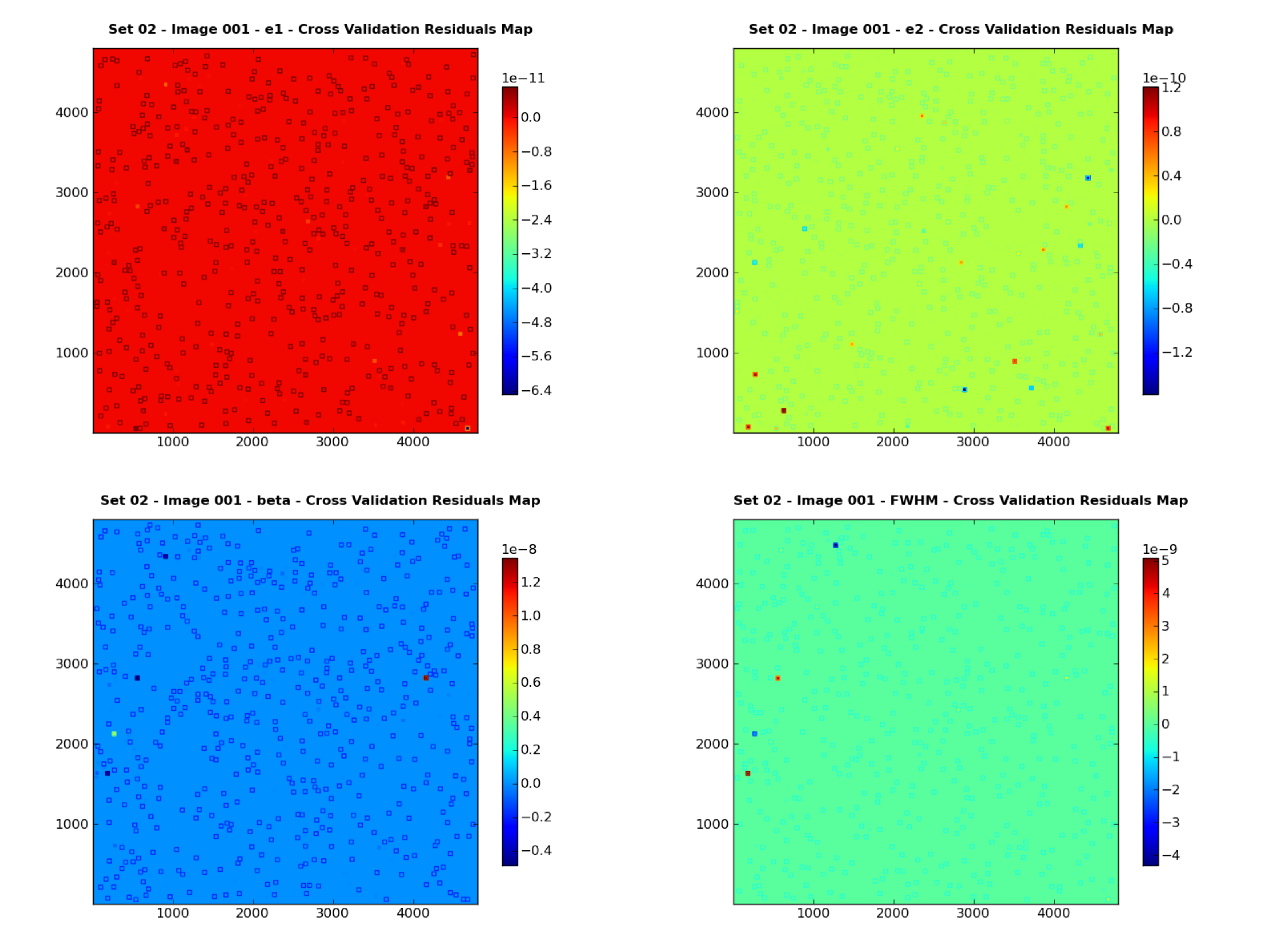








Kriging Cross-Validation



A few conclusive remarks

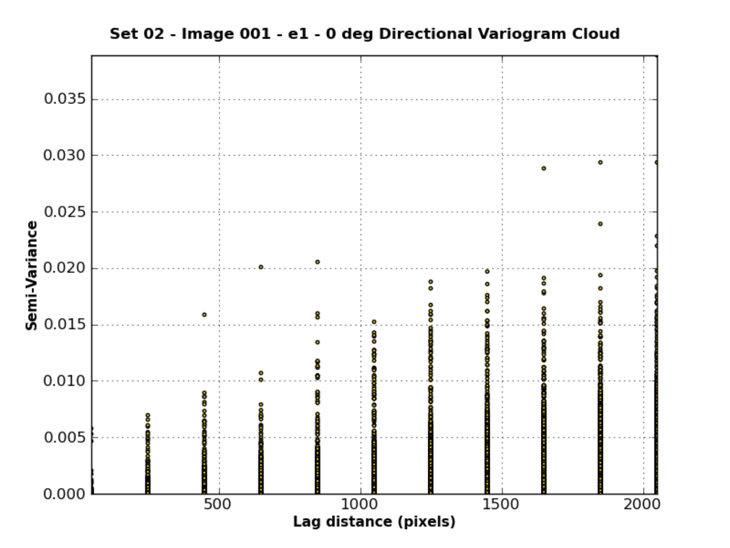
- The results of Kriging from simulated data and cross-validation are very encouraging
- Kriging seems quite sensitive to the presence of outliers or misfit values
  - The prediction currently fails for some images
- We plan to submit to GREAT10 in the next few weeks once these problems are solved
- Our current implementation of Kriging:
  - Does not take spatial anisotropy into account
  - Does not attempt to correct the trend in the data

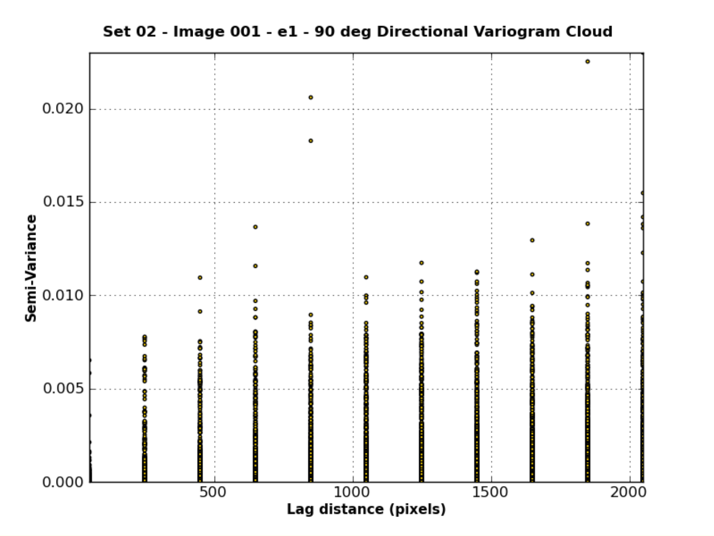
  Addressing this may improve the results further

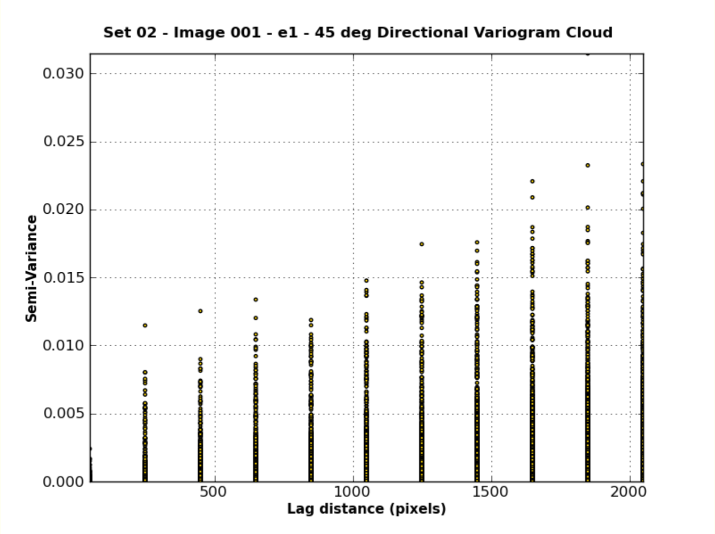
Thank You!

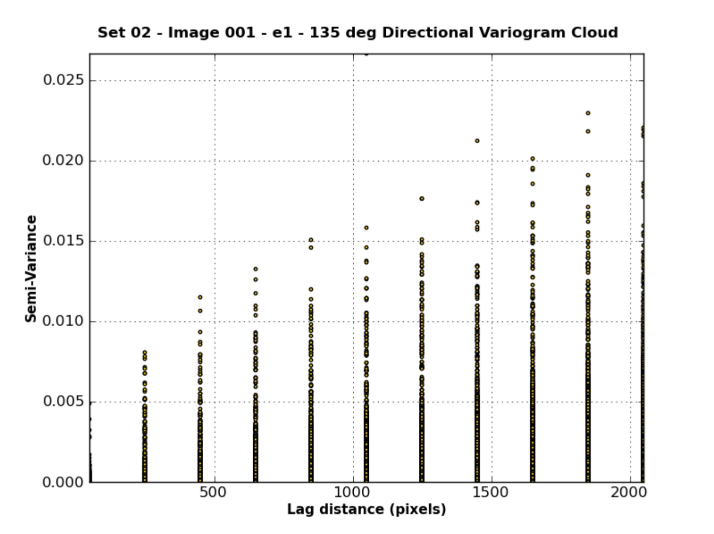
Some Extra Slides

Sample Variogram Clouds









H-Scattergram Plots