What is your confidence in your simulated confidence interval?

Danny Kentwell

SRK Consulting

srk consulting

An empirical look at simulation using skewed, non stationary data and its sensitivity to input parameters:

Difficulties working with real world gold deposits

Overview

- Abstract
- Simulation
- Confidence intervals
- Case Study
- Conclusions

Abstract

- Assessment of uncertainty in Resource estimation is often quantified by deriving a confidence interval for a particular volume from a set of block simulations.
- The underlying assumption is that the set of parameters used for input to the simulation is fixed and that they are all correct.
- Just as kriging results and kriging quality are sensitive to the number of samples in the local search neighborhood, so too are simulations that rely on local search neighborhoods for their implementation.
- For example, fewer samples in the neighborhood means a different kriging result and more importantly a larger kriging variance, this leads to a wider set of possible simulated values at each point/block and thus to a different set of confidence intervals for any given volume.

Simulation

Section One

srk consulting

Chiles and Delfiner

- A geostatistical simulation is simply a spatially consistent Monte Carlo simulation
- A geostatistical simulation does not reproduce the genetic mechanisms that generate the observed phenomenon
- Different simulation algorithms can have a greater impact on results than changes in the input parameters

What are simulations designed for

- Simulations are designed for examination of variability at a scale less than that which is directly available from the observed (sampled) information.
- To “overcome” to smoothing effects inherent on kriged estimates.
- To provide more realistic grade and tonnage curves
- To provide a range of possible estimates for project risk assessment and sensitivity analysis
- To evaluate different drill spacing and pattern configurations

Not accounted for

- Geological uncertainty
- Sampling error
- Uncertainty on the mean
- Algorithm differences
- Skewed (non Gaussian) distributions
- Non stationary domains
- Local grades
- Local changes in variography

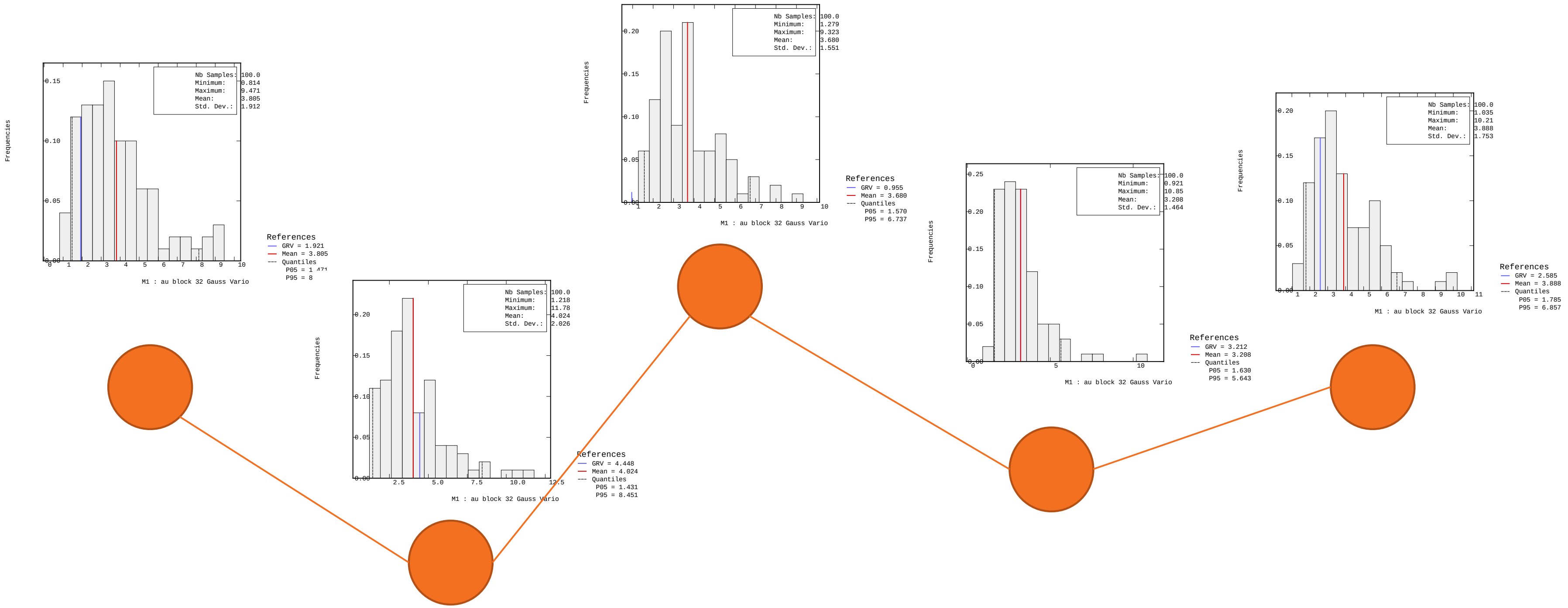
Some Simulation algorithms

- Sequential Gaussian (SK and OK)
- Turning Bands (SK and OK)
- Direct Block (SK and OK)
- Sequential Indicator
- Fractal

SGS - At each point

- Using gaussian transformed data
- Estimate (SK) mean and variance
- Randomly sample from the (Gaussian) distribution defined by the Simple Kriged mean and the Simple Kriged variance of the point
- Back transform to raw space
- Aggregate the points within a block to derive a block values

Monte Carlo


Frequencies
Nb Samples: 100.0
Minimum: 0.814
Maximum: 9.471
Mean: 3.805
Std. Dev.: 1.912
M1 : au block 32 Gauss Vario
References
GRV = 1.921
Mean = 3.805
Quantiles
P95 = 8
Frequencies
Nb Samples: 100.0
Minimum: 1.218
Maximum: 11.78
Mean: 4.024
Std. Dev.: 2.026
M1 : au block 32 Gauss Vario
References
GRV = 4.448
Mean = 4.024
Quantiles
P05 = 1.431
P95 = 8.451
Frequencies
Nb Samples: 100.0
Minimum: 1.279
Maximum: 9.323
Mean: 3.680
Std. Dev.: 1.551
M1 : au block 32 Gauss Vario
References
GRV = 0.955
Mean = 3.680
Quantiles
P05 = 1.570
P95 = 6.737
Frequencies
Nb Samples: 100.0
Minimum: 0.921
Maximum: 10.85
Mean: 3.208
Std. Dev.: 1.464
M1 : au block 32 Gauss Vario
References
GRV = 3.212
Mean = 3.208
Quantiles
P05 = 1.630
P95 = 5.643
Frequencies
Nb Samples: 100.0
Minimum: 1.035
Maximum: 10.21
Mean: 3.888
Std. Dev.: 1.753
M1 : au block 32 Gauss Vario
References
GRV = 2.585
Mean = 3.888
Quantiles
P05 = 1.785
P95 = 6.857

Outcome Dependent on

- Transformation parameters - including de-clustering weights and top cuts
- Local sample search parameters
- SK mean (in the case of sequential Gaussian simulation)
- Variogram nugget, range and anisotropy

Drawbacks of (single) simulations

- They do not provide the best local (block) estimate and are locally biased.
- The exhibit less connectivity (clustering) compared to reality
- The do not work correctly with non stationary domains (trends)
- Non Gaussian data requires a transformation and subsequent back transformation.

Reproduction of variogram

- A search neighbourhood containing too few samples can result in poor conditioning and poor reproduction of the variogram (spatial consistency)
- “Optimum” search neighbourhoods used for kriging may not be “optimum” for simulation

Confidence Intervals

Section Two

What is a Confidence Interval?

- A range of values within which we are fairly sure our true value lies

How to use 100 sims for CI?

- Classical CI on domain means
  - 95% CI = +/- (1.96 * sd/sqrt100) for assumed Gaussian errors
  - 95% CI = +/- (3* sd/sqrt100) for unknown error distribution (Chiles and Delfiner)
- Rank models by mean and examine the extremes - for example the 5^{th} percentile model and 95^{th} percentile model

CI of what?

- Individual point mean grades?
- Individual block mean grades? - Local
- Entire domain mean grade? - Global
- Entire domain grade at a specified cut off?
- Quarterly or annual production tonnage at a specified cut off?
- Quarterly or annual production grade at a specified cut off?
- Quarterly or annual production metal at a specified cut off?

Case Study

Section Three

Case study

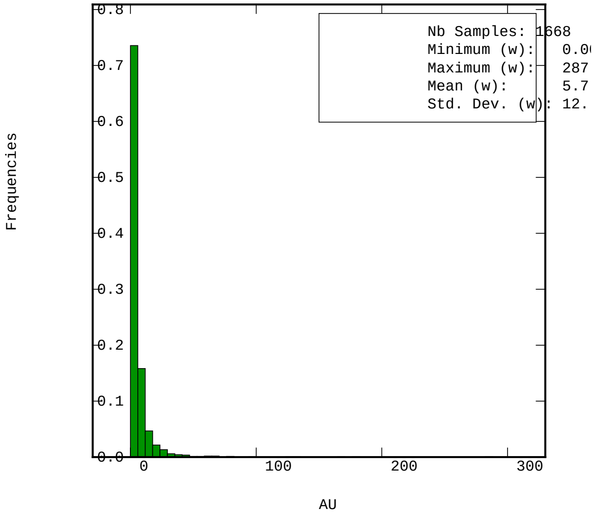
- Shear hosted gold
- Highly skewed
- Non stationary

- OK with increasing sample numbers in search neighborhood
- SGS (SK) with increasing sample numbers in search neighborhood

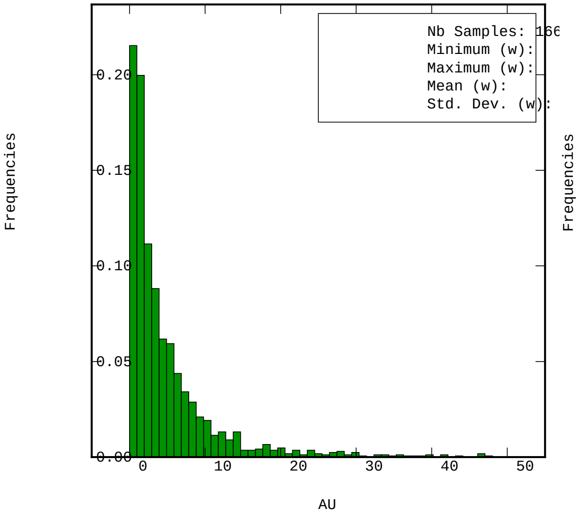
- 8 32 and 128 sample neighborhoods

- 32 being “optimal” for OK (maximize regression slope and minimize negative weights)

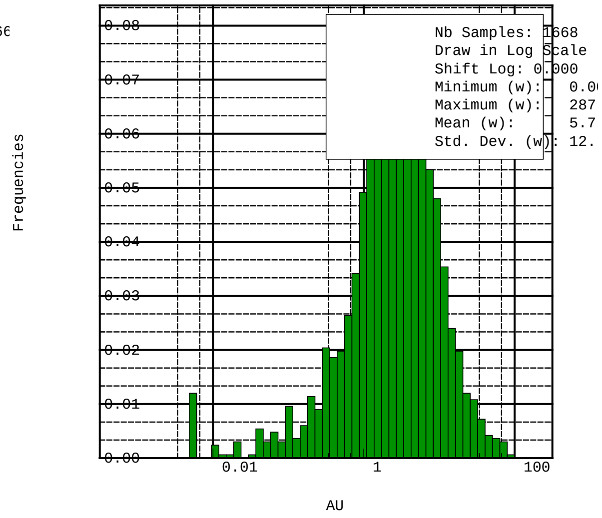
Sample Data



Nb Samples: 1668
Minimum (w): 0.0
Maximum (w): 287
Mean (w): 5.7
Std. Dev. (w): 12.
Frequencies
AU

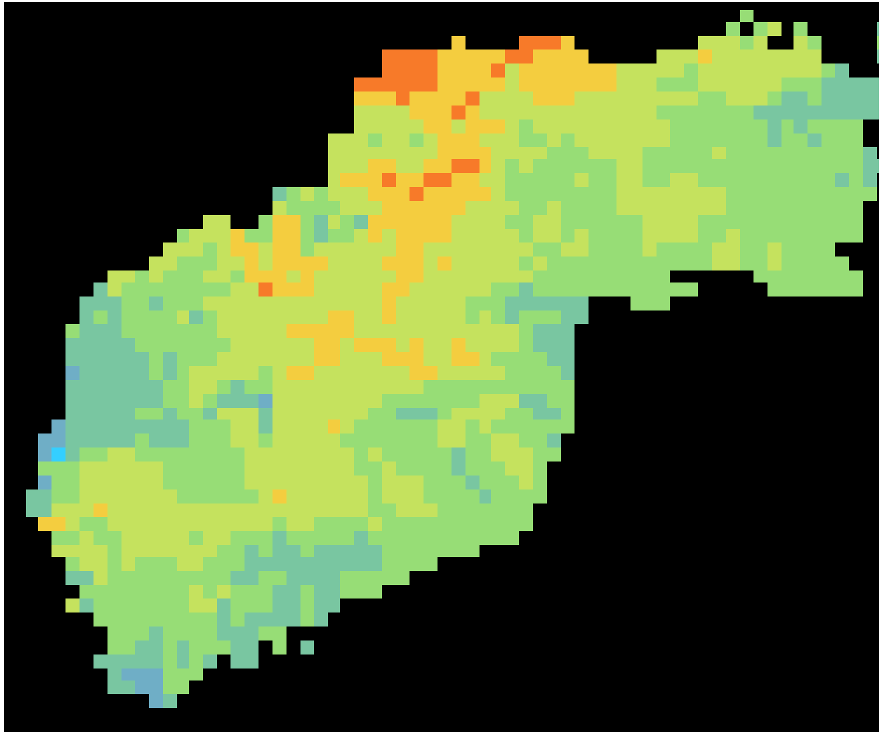


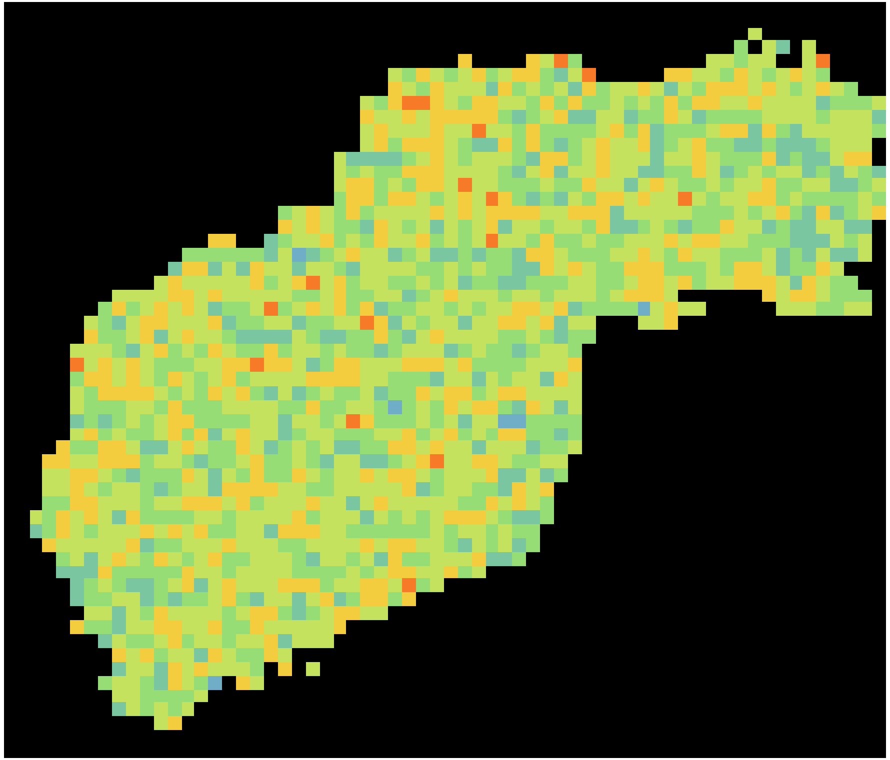
Nb Samples: 166
Minimum (w):
Maximum (w):
Mean (w):
Std. Dev. (w):
Frequencies
AU



Nb Samples: 1668
Draw in Log Scale
Shift Log: 0.000
Minimum (w): 0.0
Maximum (w): 287
Mean (w): 5.7
Std. Dev. (w): 12.
Frequencies
AU

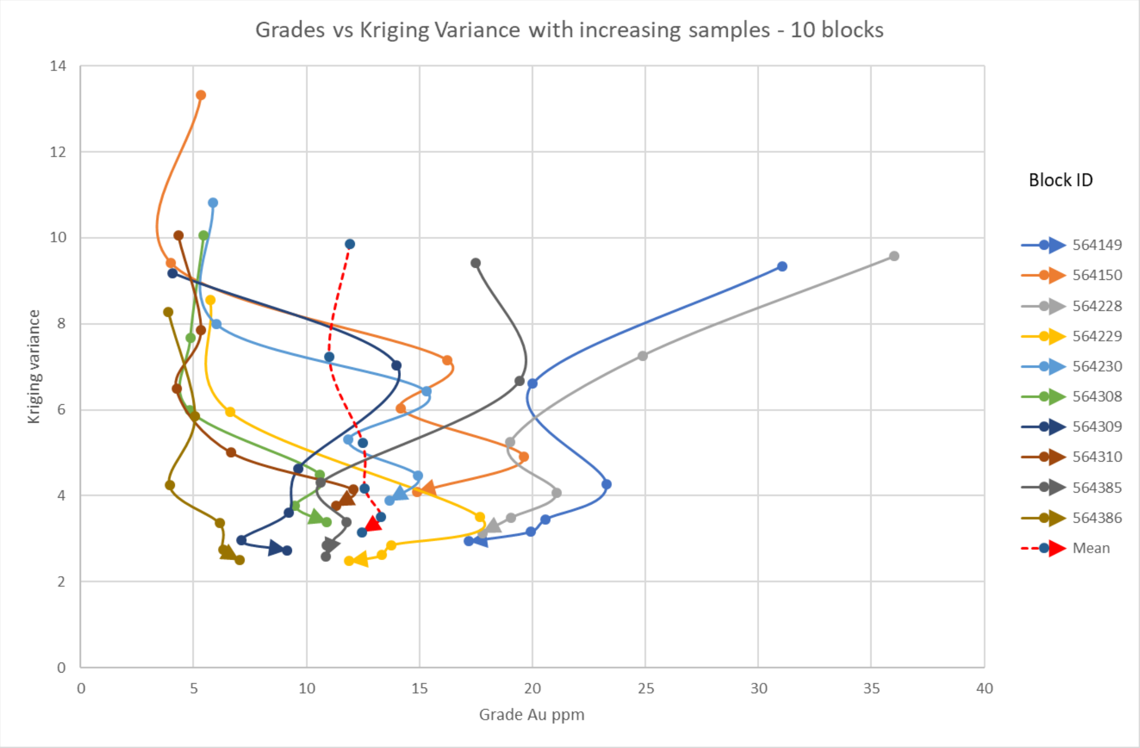
Estimation (OK 32) vs Simulation (SGS SK 32)



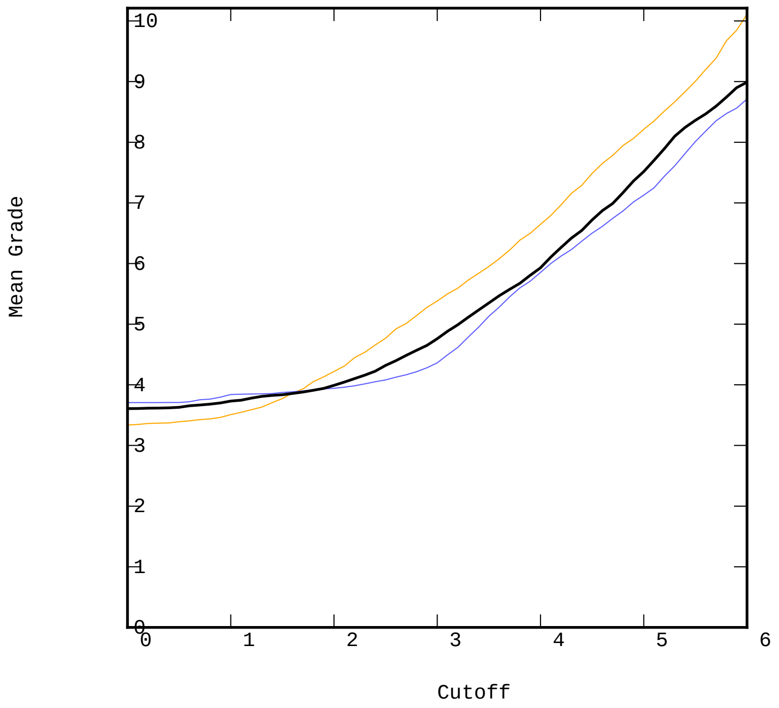


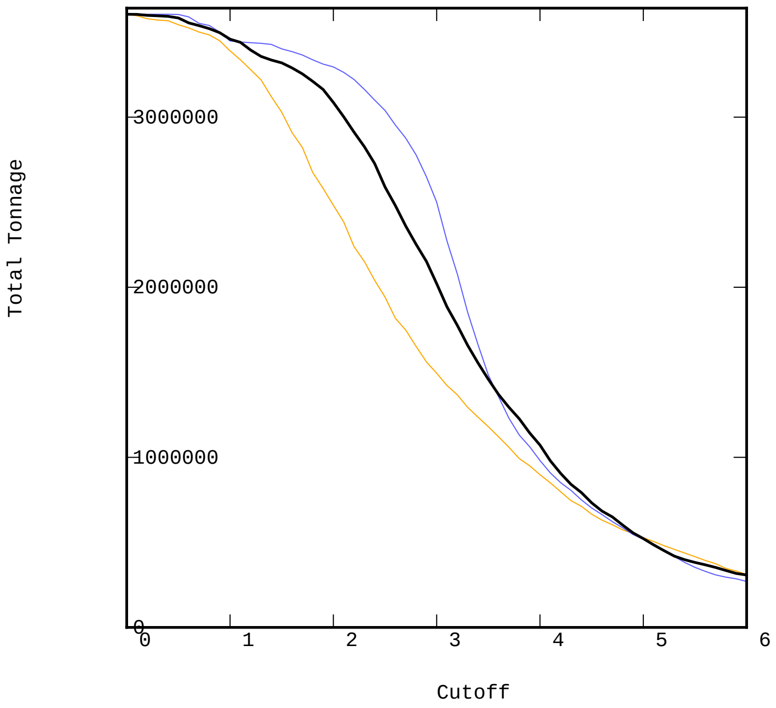
Variance Reduction with increasing samples (OK estimate)

4
8
16
32
64
128

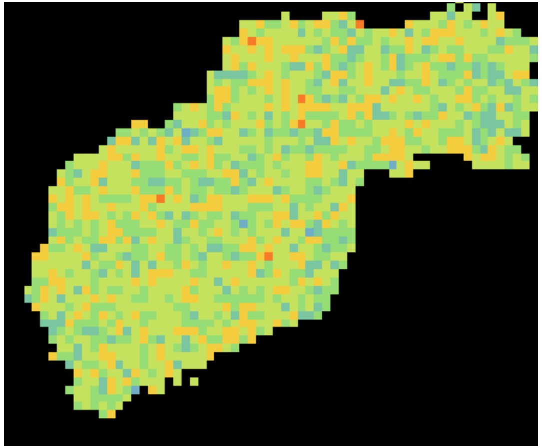


OK Grade and Tonnage curves 8 32 128

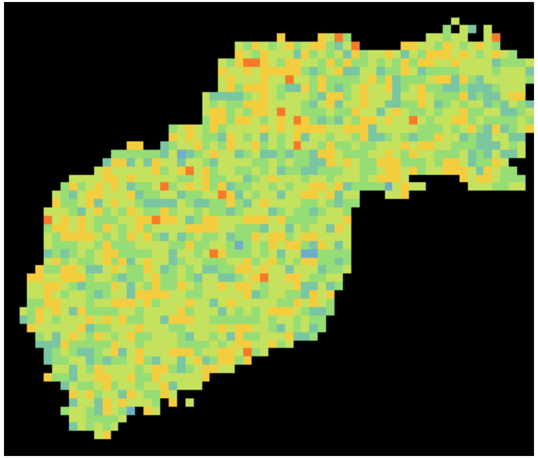




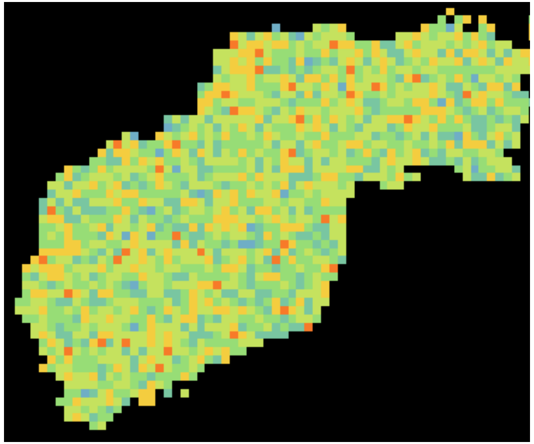
SGS (SK) Single Sims



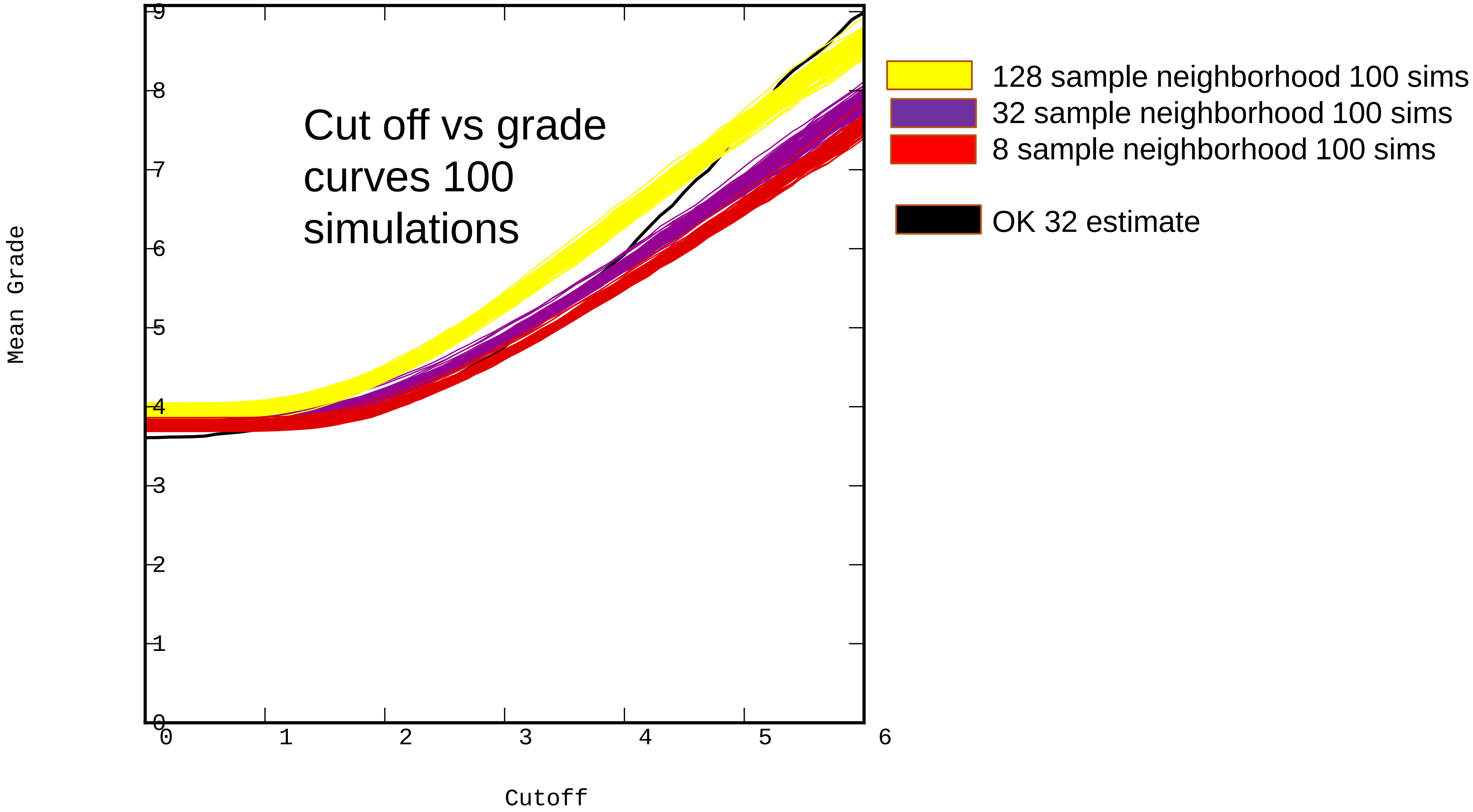
8 samples



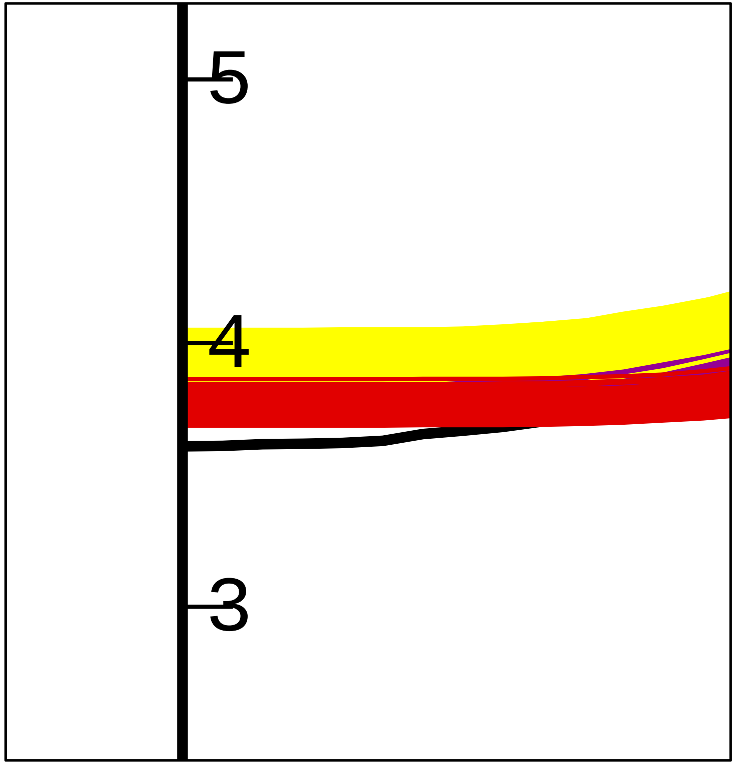
32 samples

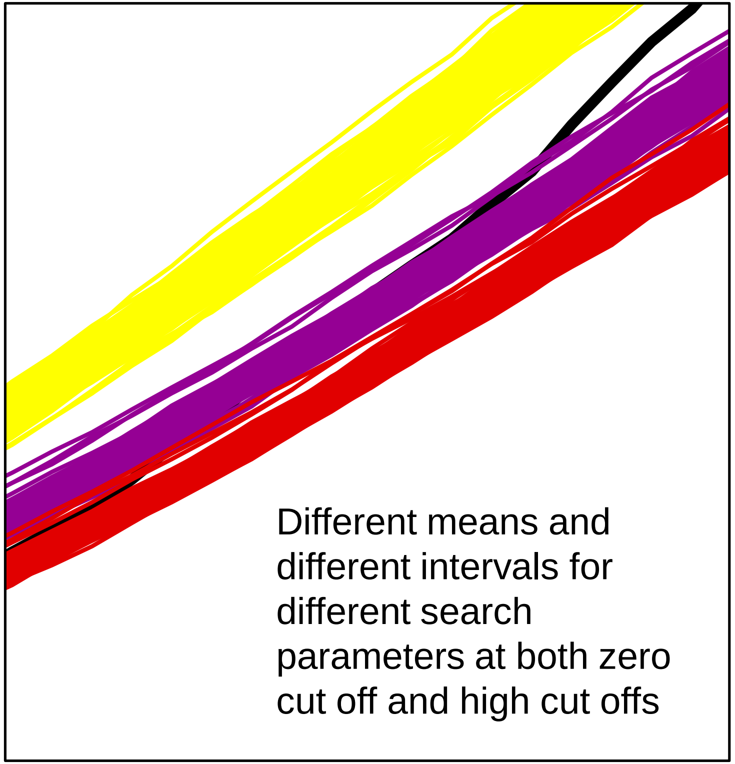


128 samples


Cut off vs grade
curves 100
simulations
128 sample neighborhood 100 sims
32 sample neighborhood 100 sims
8 sample neighborhood 100 sims
OK 32 estimate
Mean Grade
Cutoff
0
1
2
3
4
5
6
7
8
9
0
1
2
3
4
5
6

What set to use for risk modelling?





Practical issues with simulation (SGS and TB)

- Need to transform to and from Gaussian
  - Complicated by possible top cut application before transformation
  - Complicated by the need to de-cluster before transformation
  - (Both of these may be required to better approximate the true mean)
- Requirement to work with SK and stationary domains for correct results – (real world gold deposit domains are almost never stationary).

Conclusions 1

- Conventional geostatistical simulations ………

- Only represent the variability in estimation from a specific algorithm and a specific set of parameters
- Can under represent the likely extremes
- Can mis-represent the absolute values of average grades above a cut off

- Specifically, simulation ranges, particularly at higher cut offs can be highly sensitive to search neighborhood parameters

Some examples of other work

- Incorporation of uncertainty on input parameters (Maximum Likelihood – Dowd & Pardo-Iguzquiza 2002)

- Incorporation of geological simulation of domains prior to simulation of grade (Jewbali, Perry, Allen and Inglis 2014)

Thank You

Questions?

Simulated distribution - one block