Mineral Systems as Chemical Reactors with no Mathematics

Alison Ord and Bruce Hobbs

Mineral Dynamics
Fremantle

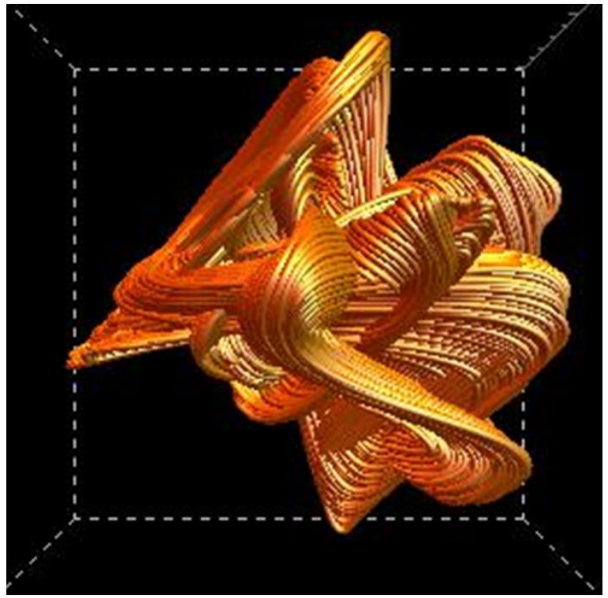
Bruce Hobbs: bruce.hobbs@csiro.au
Alison Ord: aliup.oz@gmail.com

Session 2 09.15 to 09.45.

Data and what does it mean?

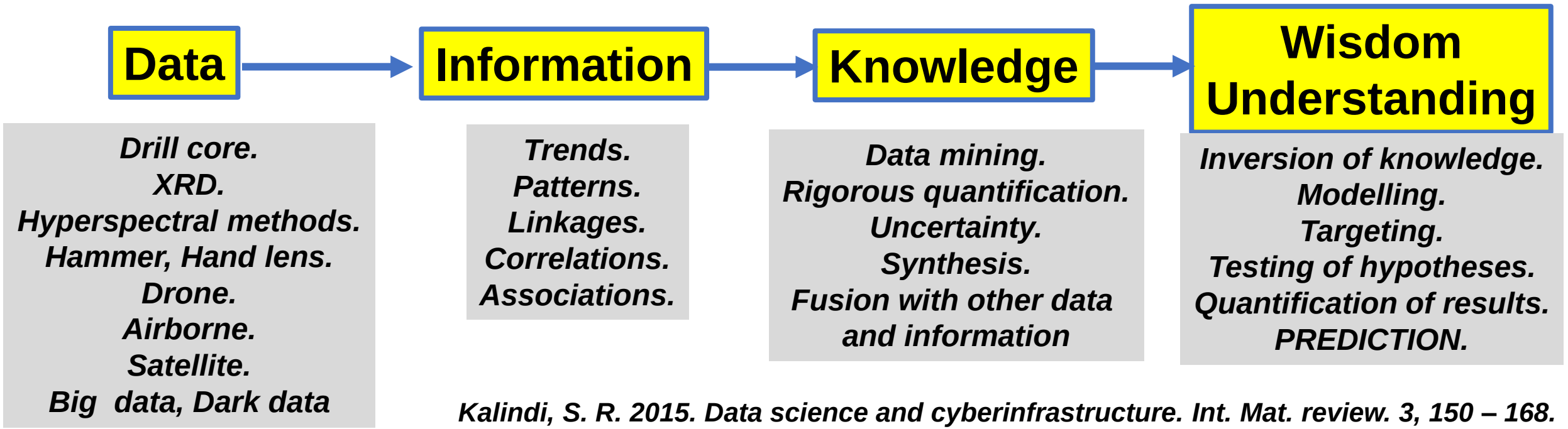
How can I use it?

What do I do with all these data?



Session 1.

Data and what does it mean? How can I use it? What do I do with all these data?



Kalindi, S. R. 2015. Data science and cyberinfrastructure. Int. Mat. review. 3, 150 – 168.

Dark data
From Wikipedia.

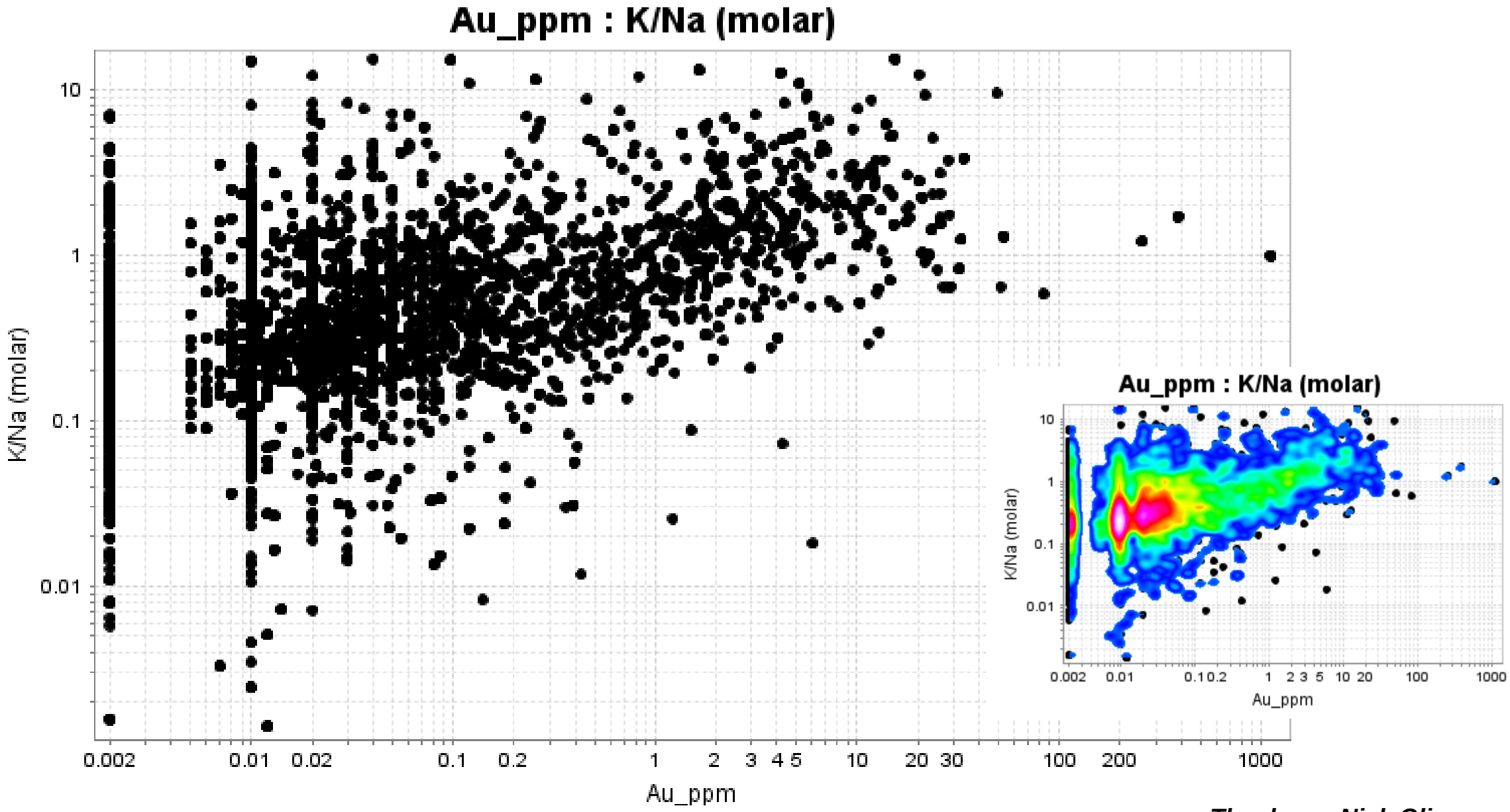
Dark data is data which is acquired through various computer network operations but not used in any manner to derive insights or for decision making.

The ability of an organisation to collect data can exceed the throughput at which it can analyse the data.

In some cases, the organisation may not even be aware that the data is being collected.

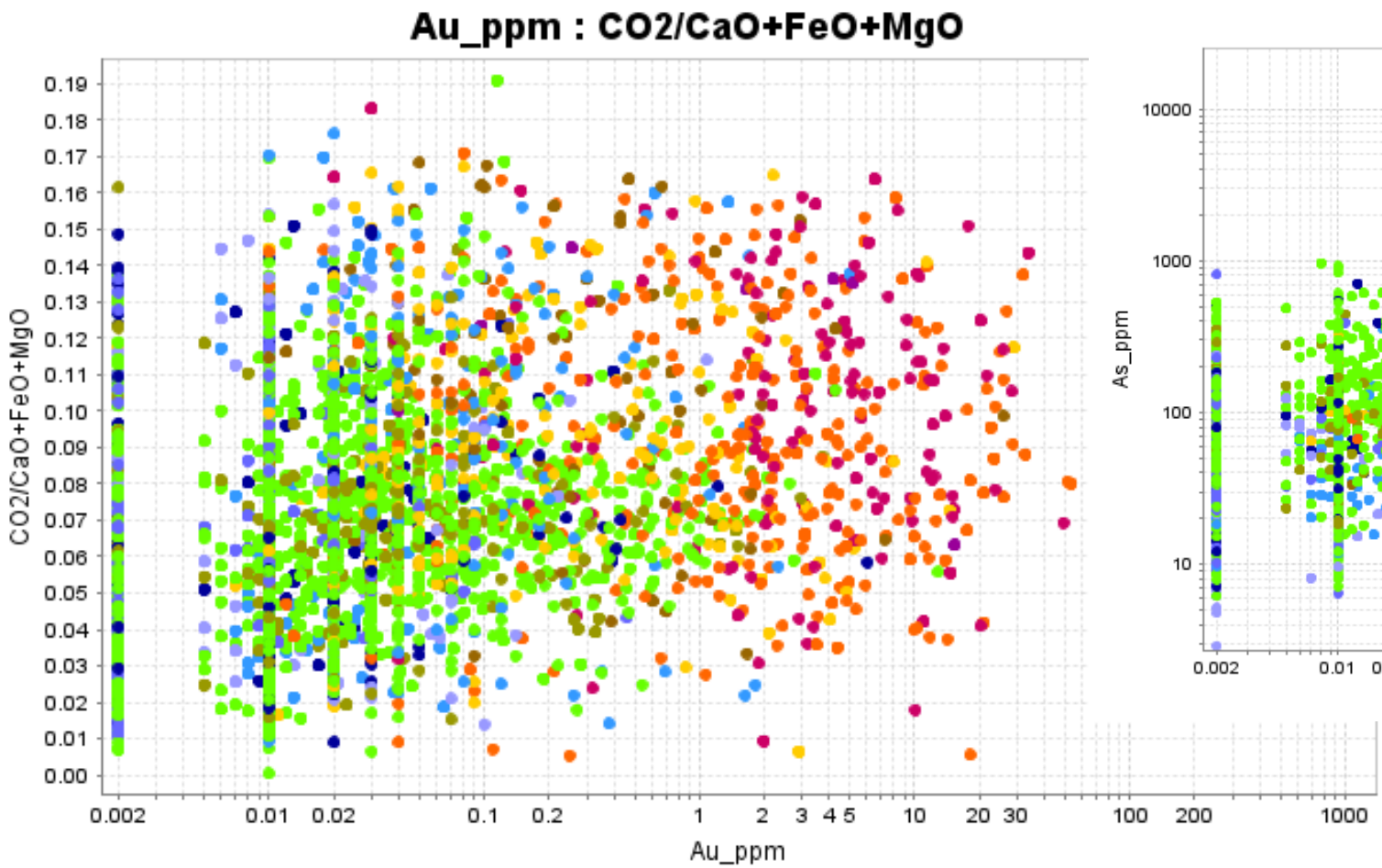
IBM estimate that roughly 90 percent of data generated by sensors and analog-to-digital conversions never get used.

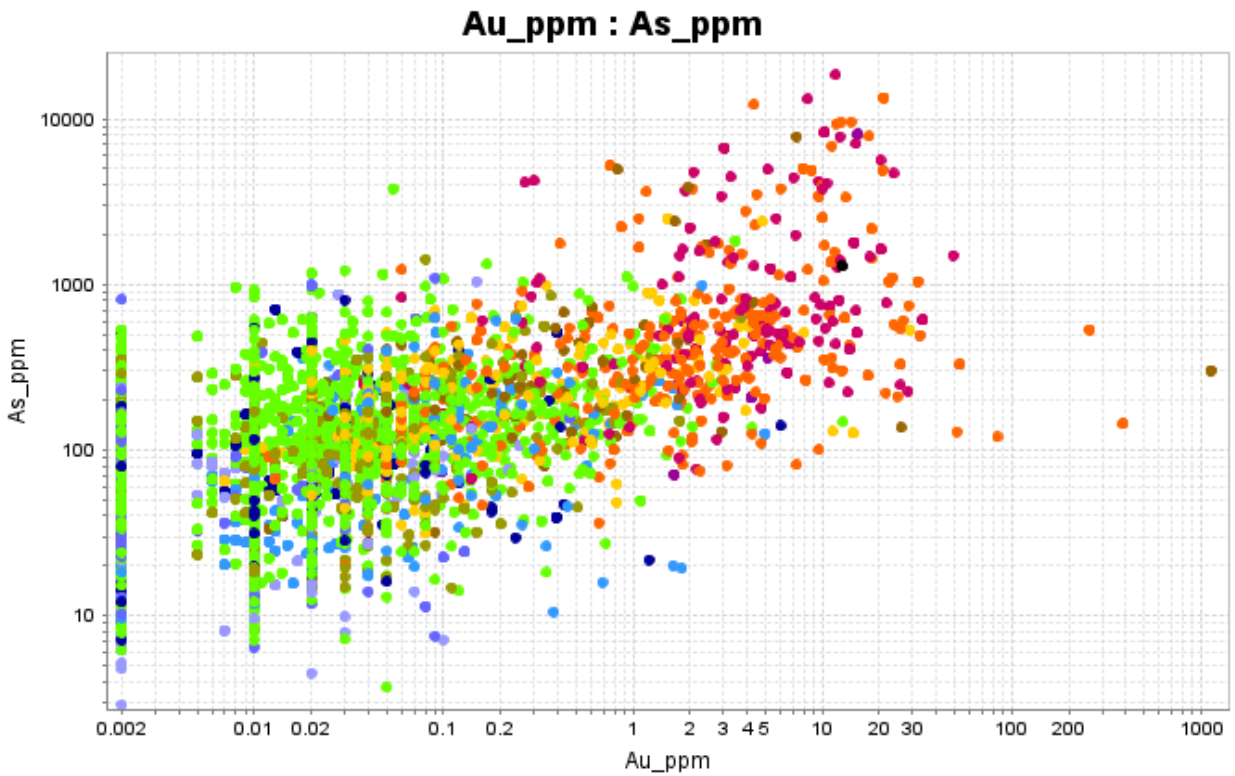
Attempts at correlations



Thank you Nick Oliver

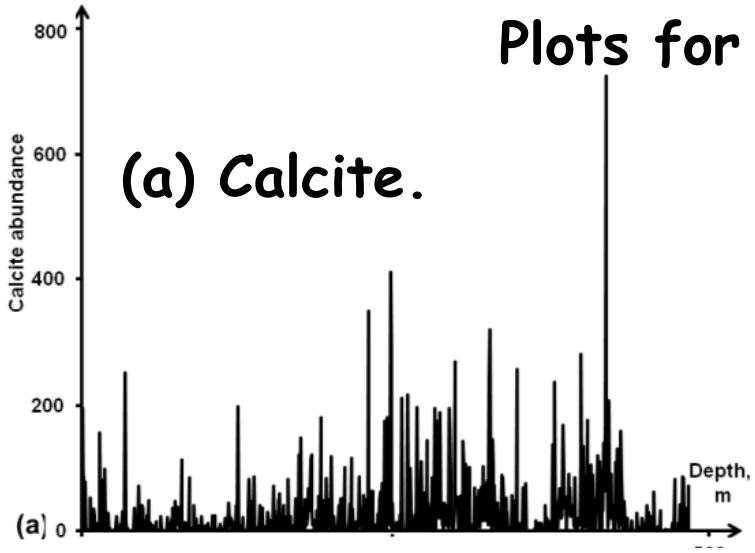
Attempts at correlations

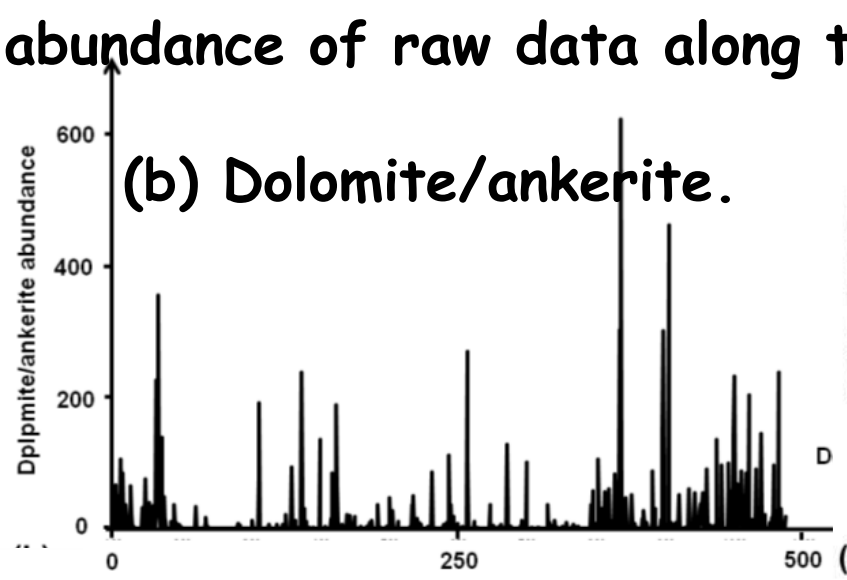


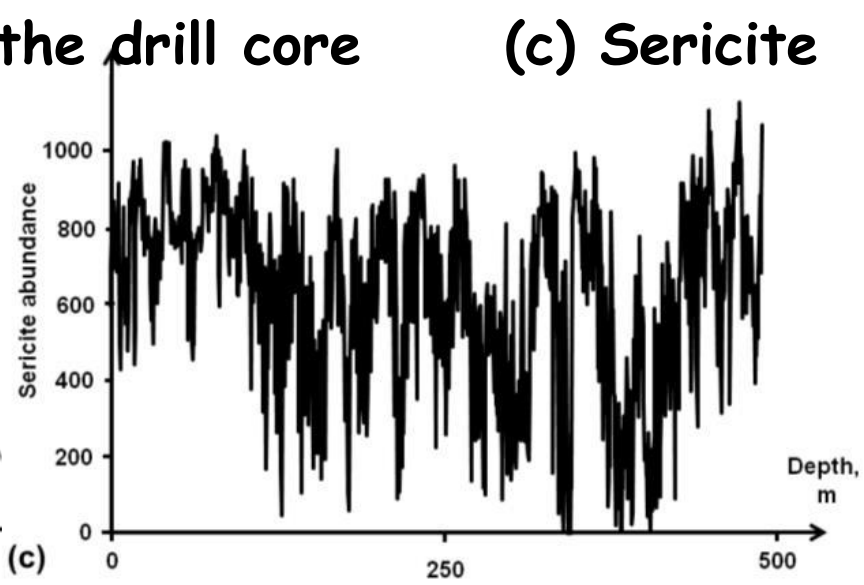


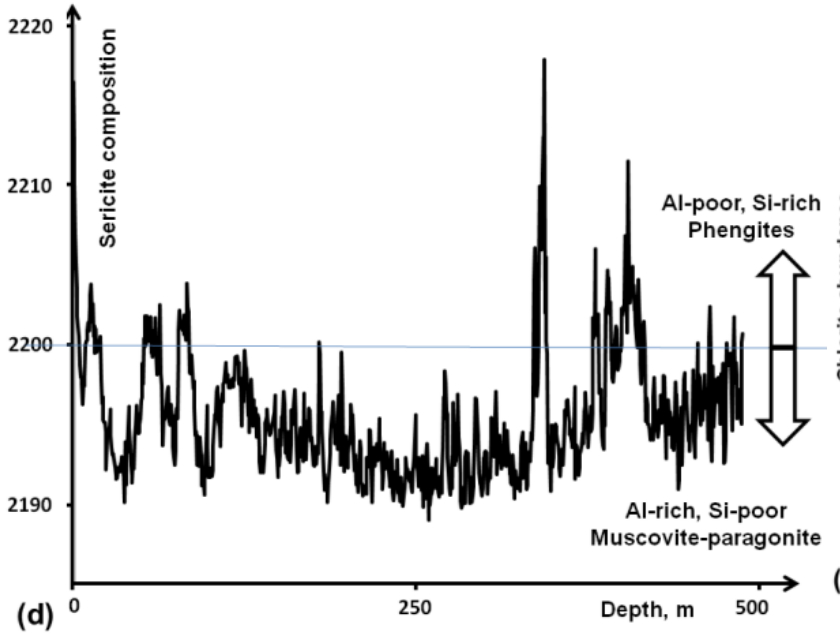
Thank you Nick Oliver

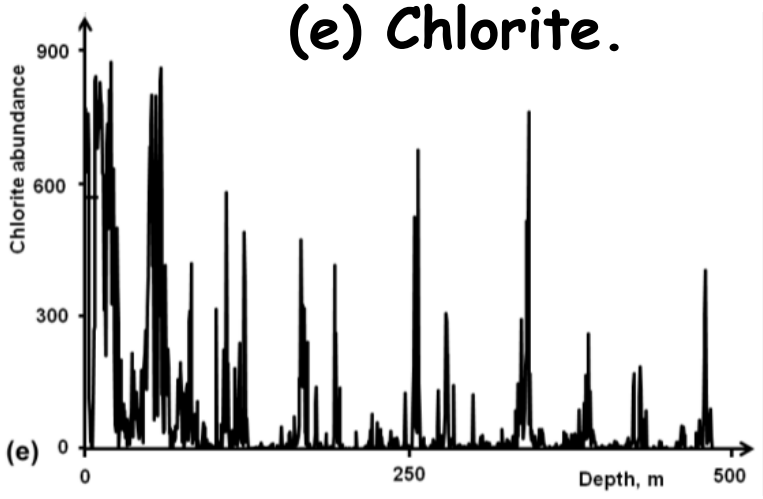
Plots for abundance of raw data along the drill core

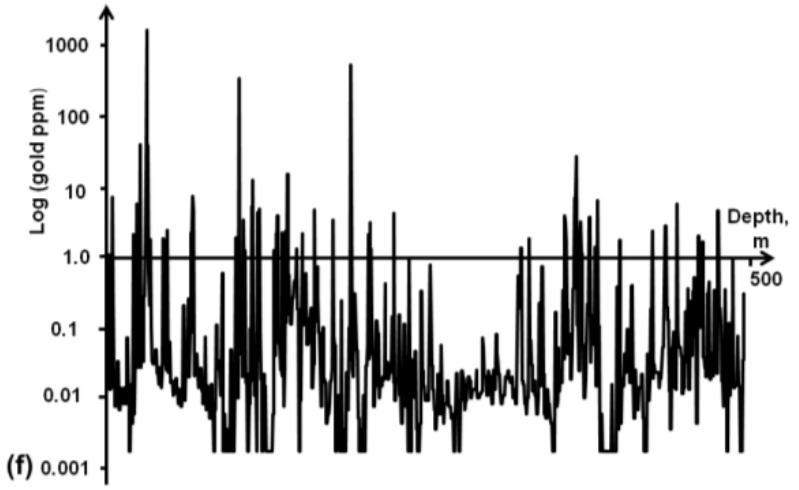








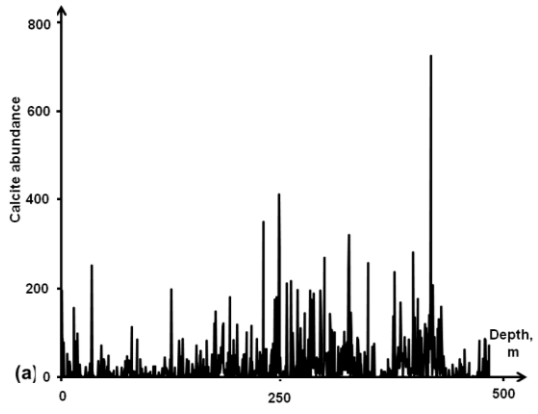


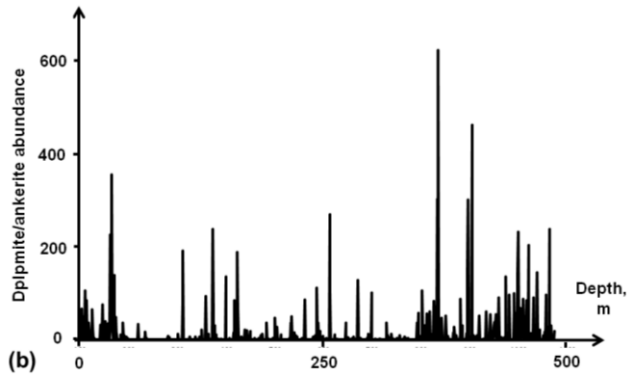


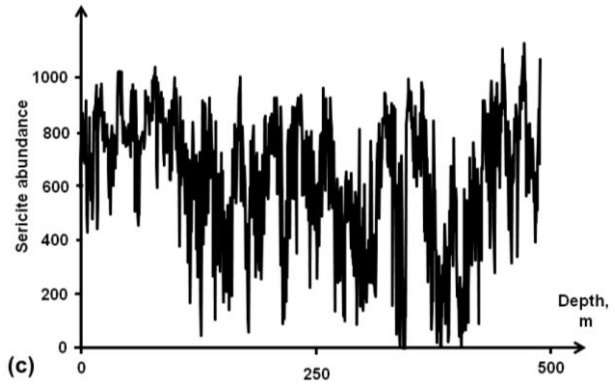
(d) Sericite composition. Vertical axis is IR absorption wavelength in nano-metres.

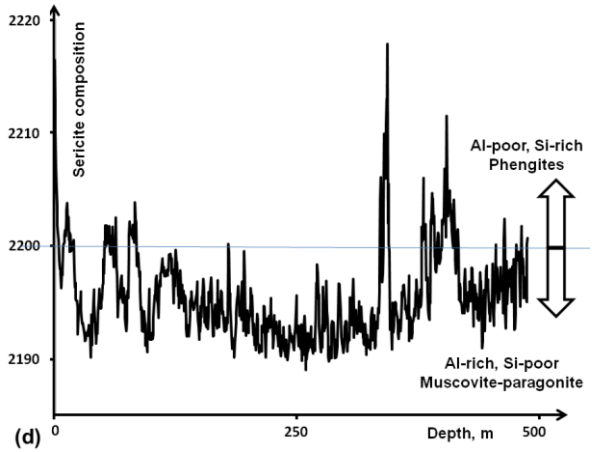
(f) Logarithm of gold concentration in ppm by weight.

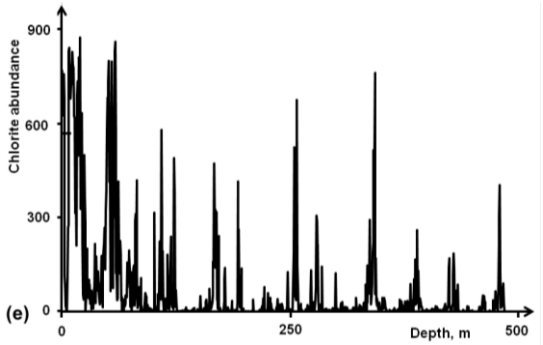
Data from single drill hole in Yilgarn Gold Deposit

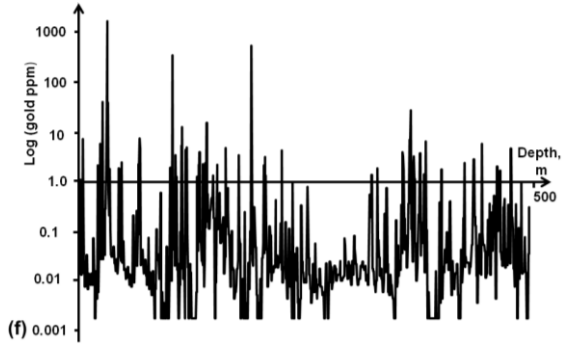


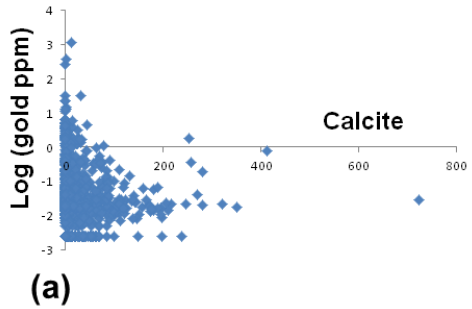


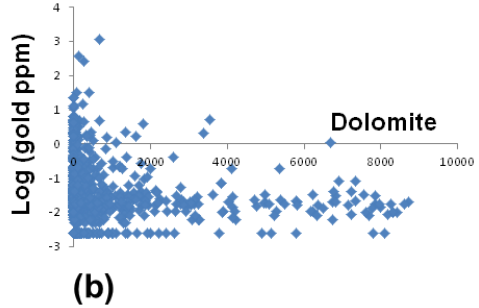


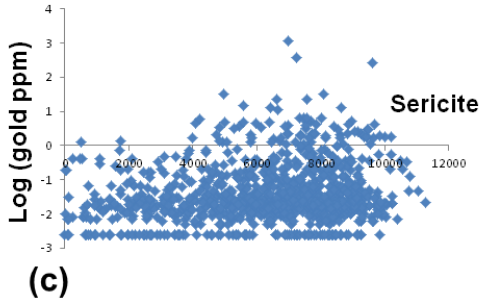


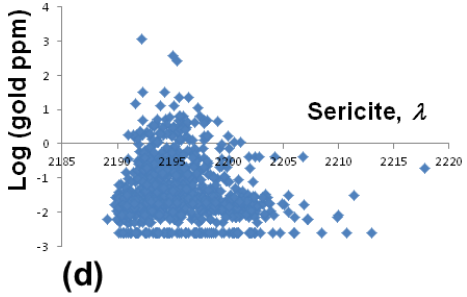












Cross correlation between signals.

	Dolomite	Log-gold	Sericite	Calcite	Chlorite	Sericite composition
Dolomite	1					
Log-gold	0.193	1				
Sericite	0.018	0.169	1			
Calcite	0.026	-0.088	-0.082	1		
Chlorite	-0.039	0.013	0.138	-0.060	1	
Sericite composition	-0.022	-0.096	-0.228	0.064	-0.024	1

Drill core.
XRD.
Hyperspectral methods.
Hammer, Hand lens.
Drone.
Airborne.
Satellite.
Big data, Dark data

Trends.
Patterns.
Linkages.
Correlations.
Associations.

Data mining.
Rigorous quantification.
Uncertainty.
Synthesis.
Fusion with other data and information

In general correlations are weak or statistically non-existent (in the classical sense).

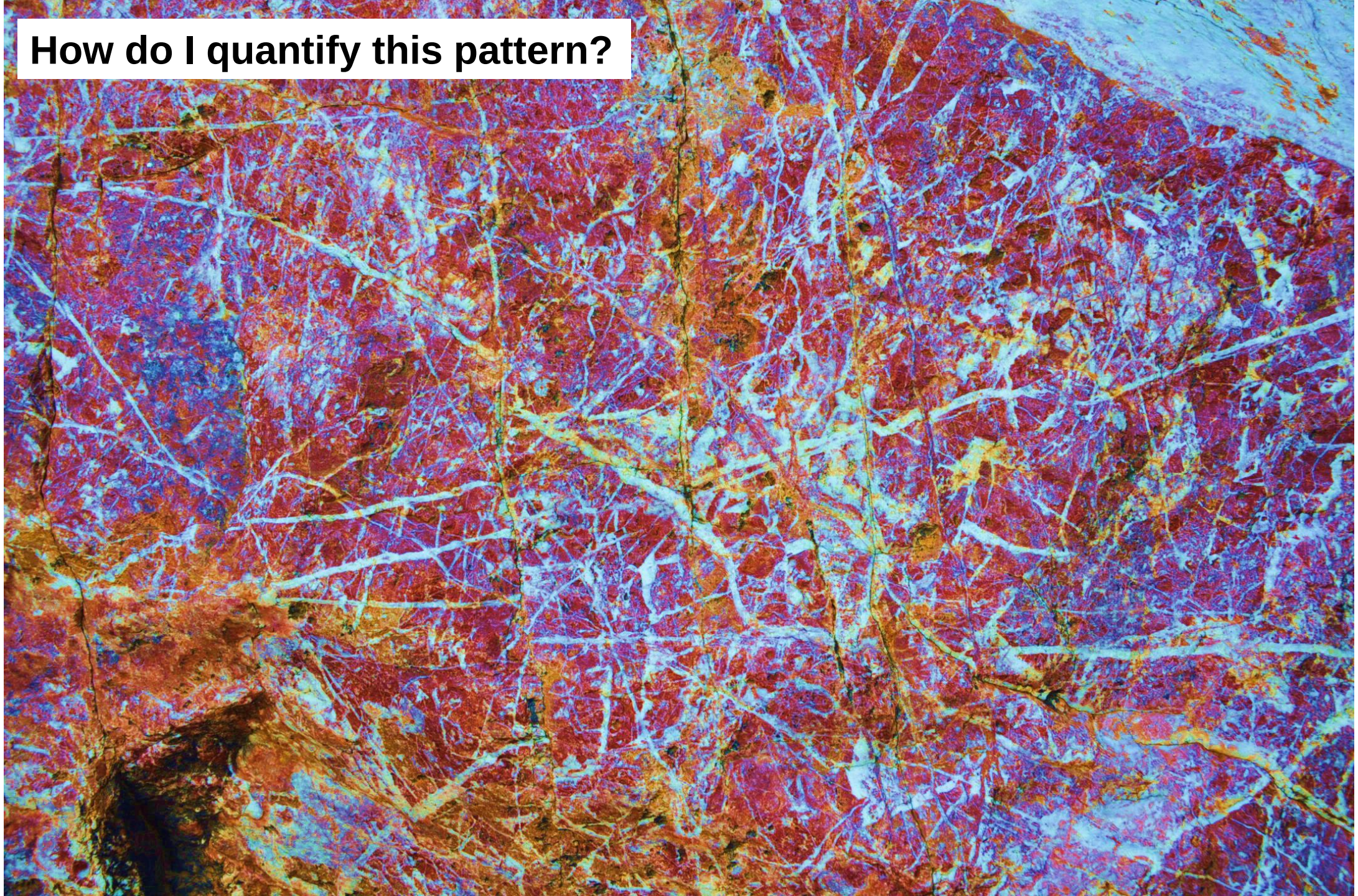
Linear correlations are always sought.

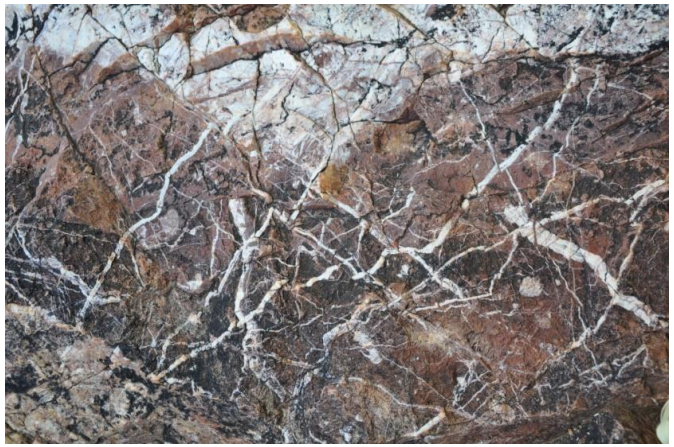
Trends, Patterns, Linkages, Correlations, Associations use classical parametric, linear methods with passing use of quantitative measures

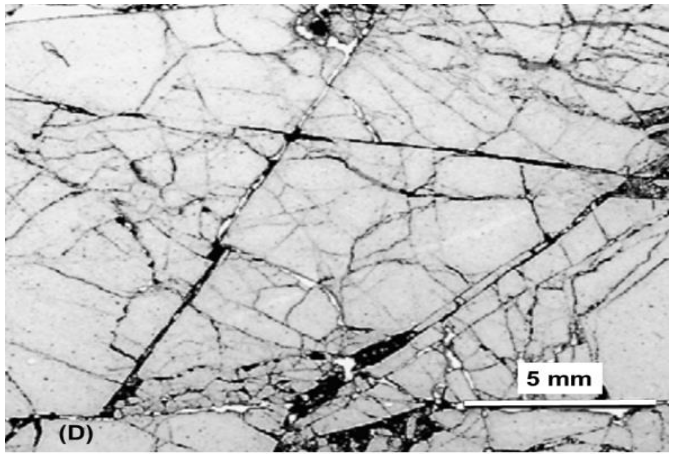
Modern methods of data mining not used. Non parametric approaches rare.

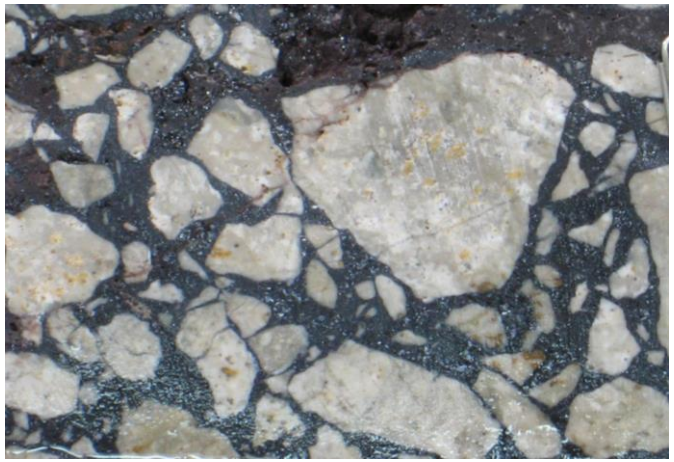
Big data becoming the norm but most remains dark data (no light is shone on the problem)


How do I quantify this pattern?


How do I quantify this pattern?

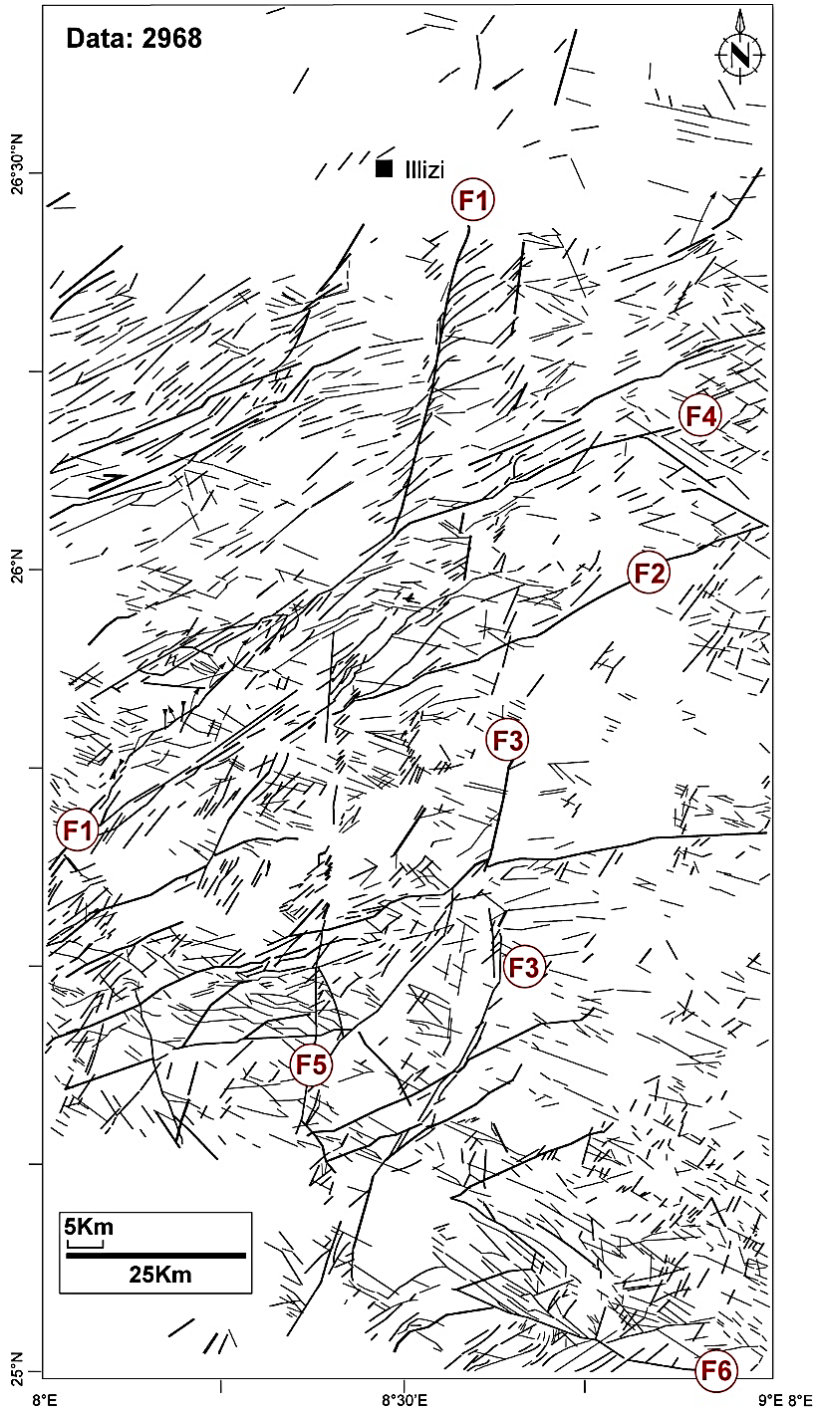






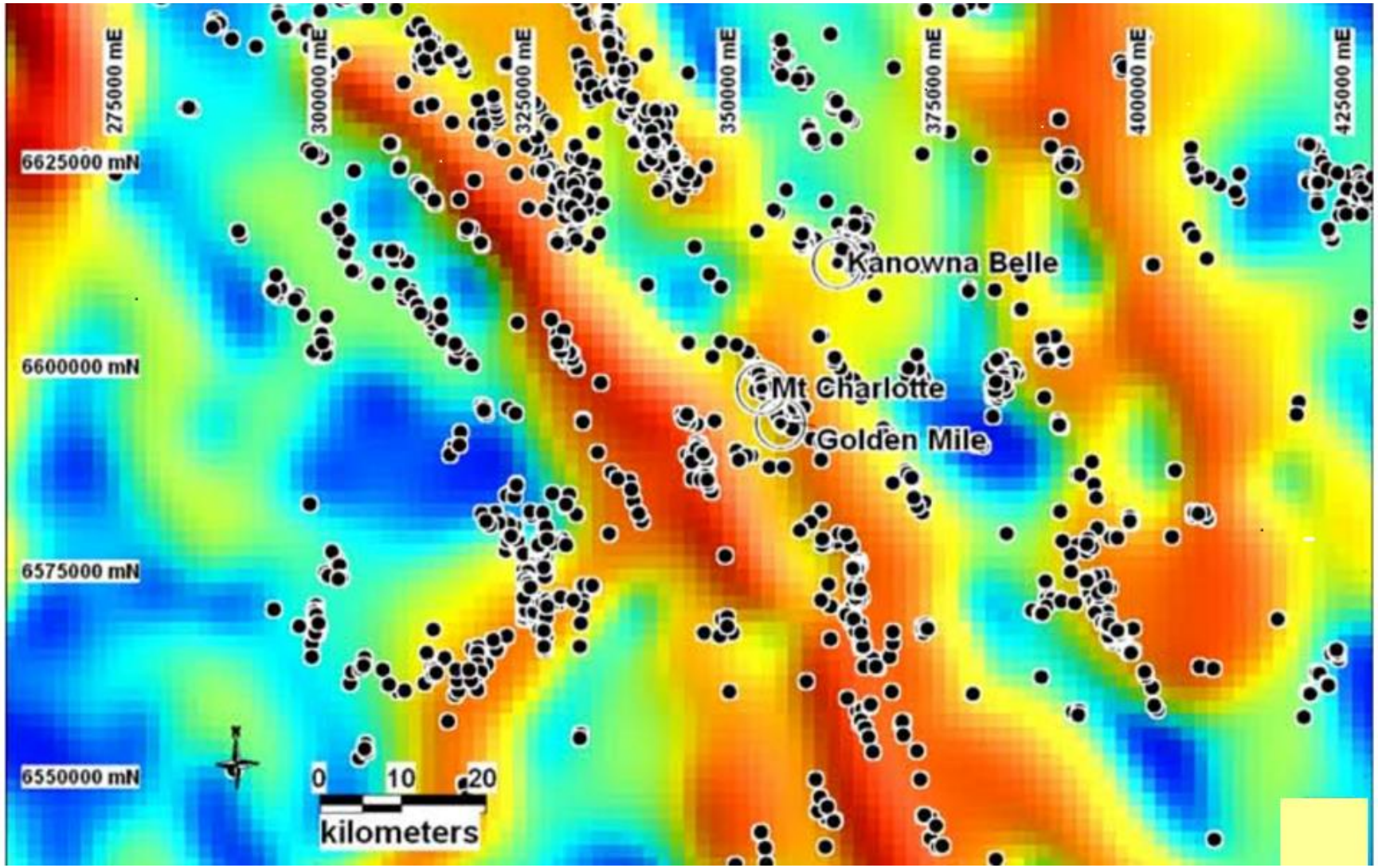
How do we quantify the nature and the differences between these fabrics in a rheologically significant manner?

How do I quantify this pattern?



Fracture system Iran

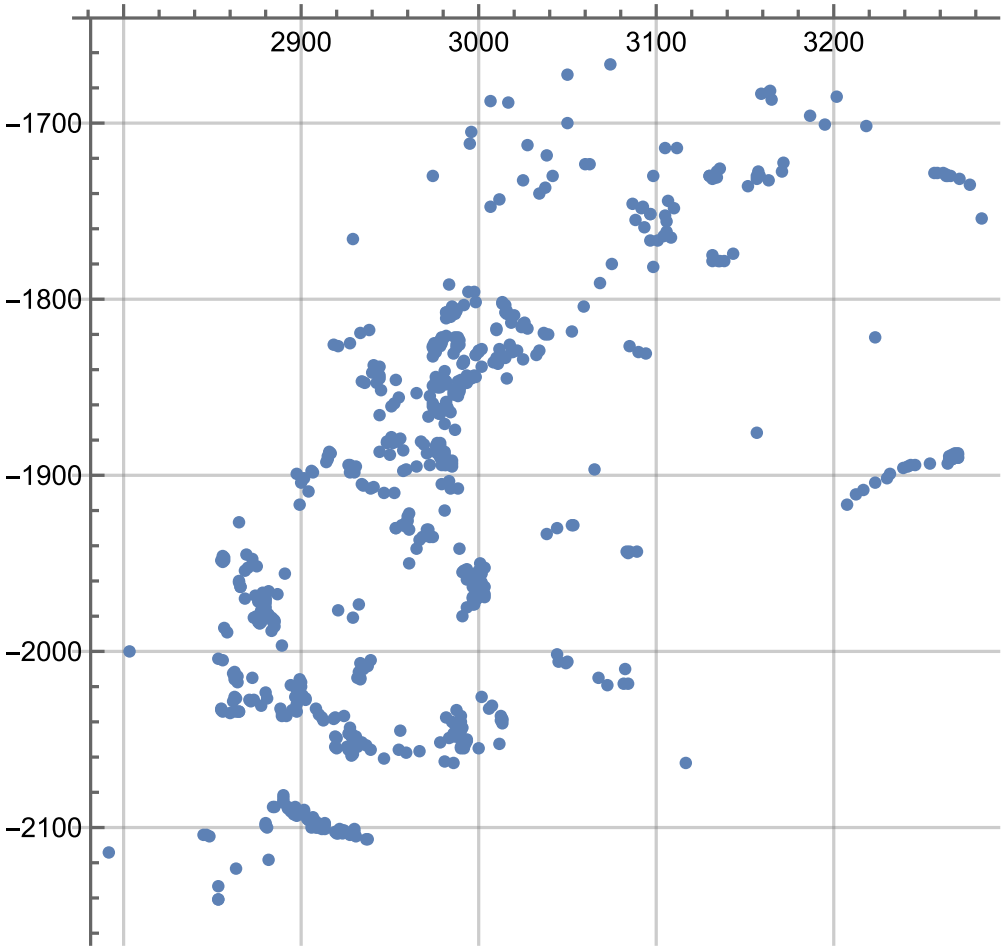
How do I quantify this pattern?



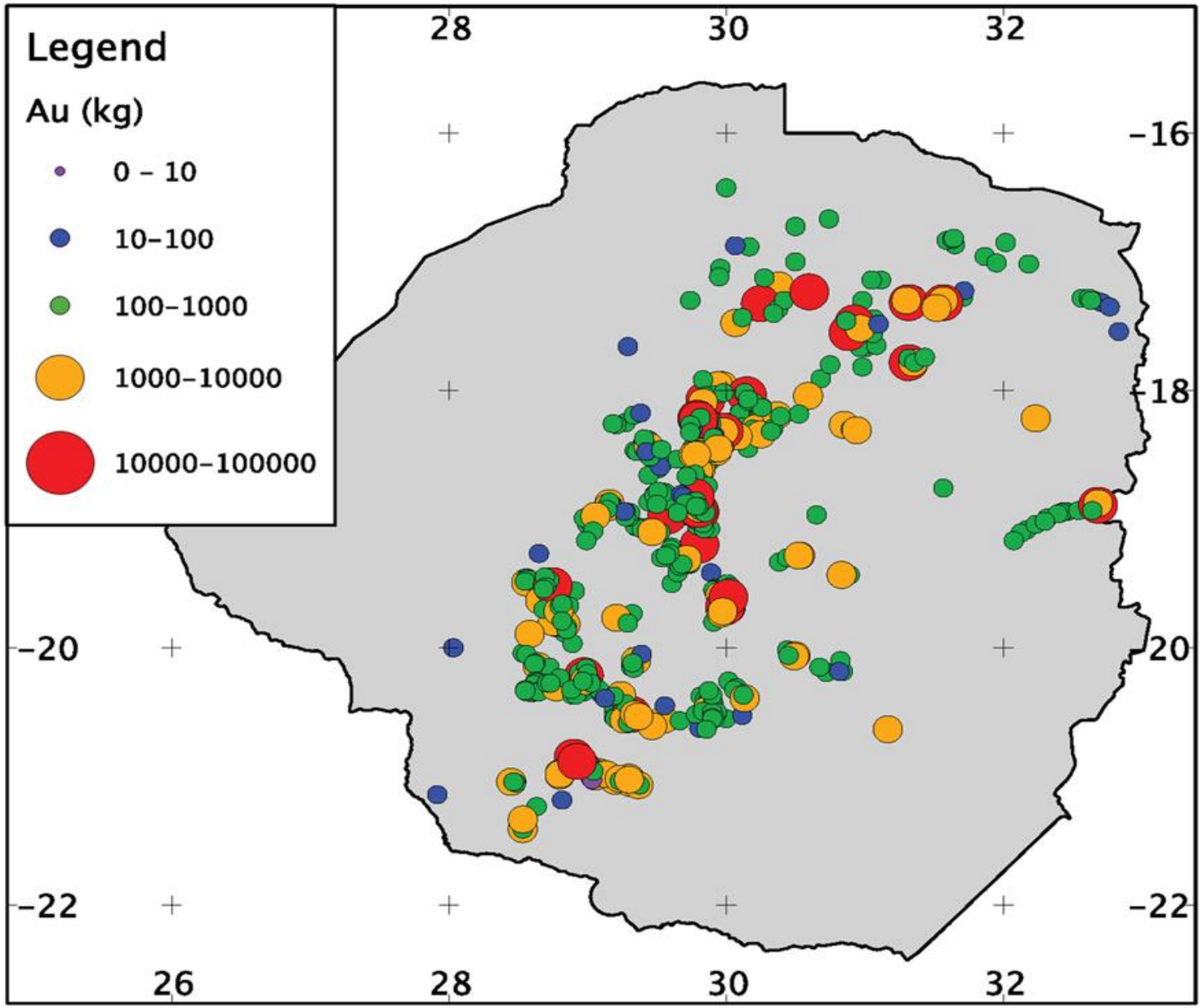
Is there anything special about this pattern?

Can I use the observed spatial distribution to target a new site?

How do I quantify this pattern?



Is there some metric that distinguishes this pattern from the Yilgarn pattern?



Zimbabwe Gold Deposits

Tools available to analyse and understand nonlinear systems.

Parametric linear approaches	Non-parametric nonlinear approaches
Mean, standard deviation	Heavy tailed distributions, differential entropy
Weights of evidence	Wavelet analysis
Moving averages	Hurst exponents
Kriging and co-kriging, covariance	Multifractals
Linear regression	Recurrence plots
Auto-correlation	Recurrence quantification measures
Fourier analysis	Nonlinear noise reduction. Nonlinear prediction. Lyapunov exponent.
Wavelet analysis	Singular value decomposition; Hankel matrix; attractor construction
Machine learning **Pattern recognition** **Clustering** **Singular value decomposition** **Detection and classification of boundaries**	Machine learning Data driven searches of equation libraries Singular value and dynamic mode decomposition Physics based model discovery Data driven prediction
	Network analysis; adjacency matrix
	N-point correlation functions
	Statistical reconstruction of structure/fabric

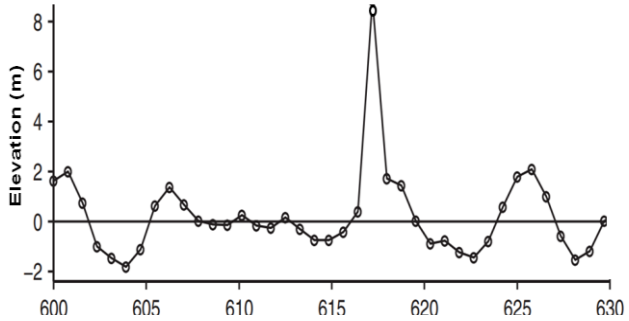
A parametric approach assumes a probability distribution such as Gaussian or log-normal and the subsequent analysis assumes the data fit that distribution. A non-parametric approach makes no such assumption and "lets the data speak for themselves".

Are Giant Ore Deposits Rogue Waves or Dragon-Kings?

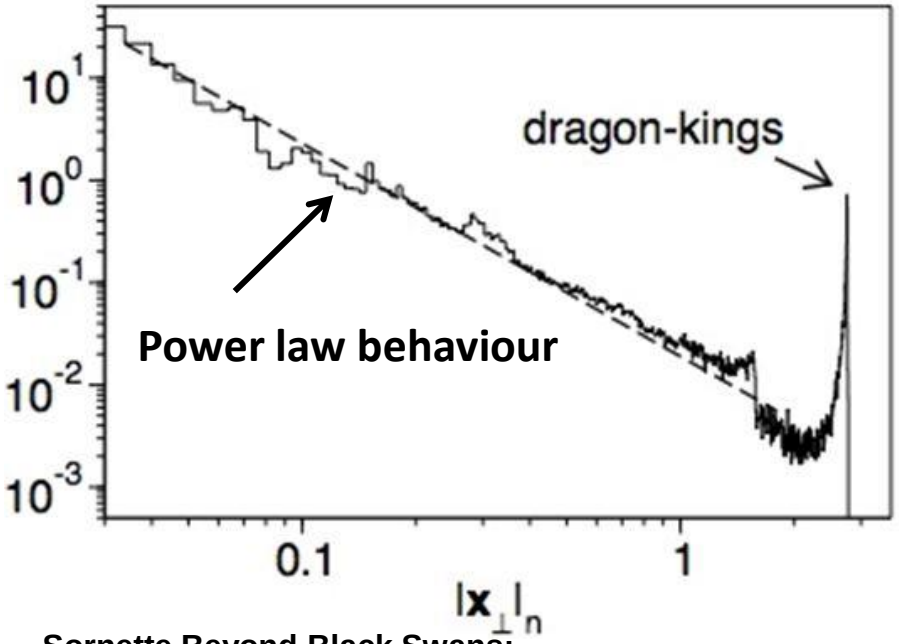
Bruce Hobbs: bruce.hobbs@csiro.au
Alison Ord: aliup.oz@gmail.com

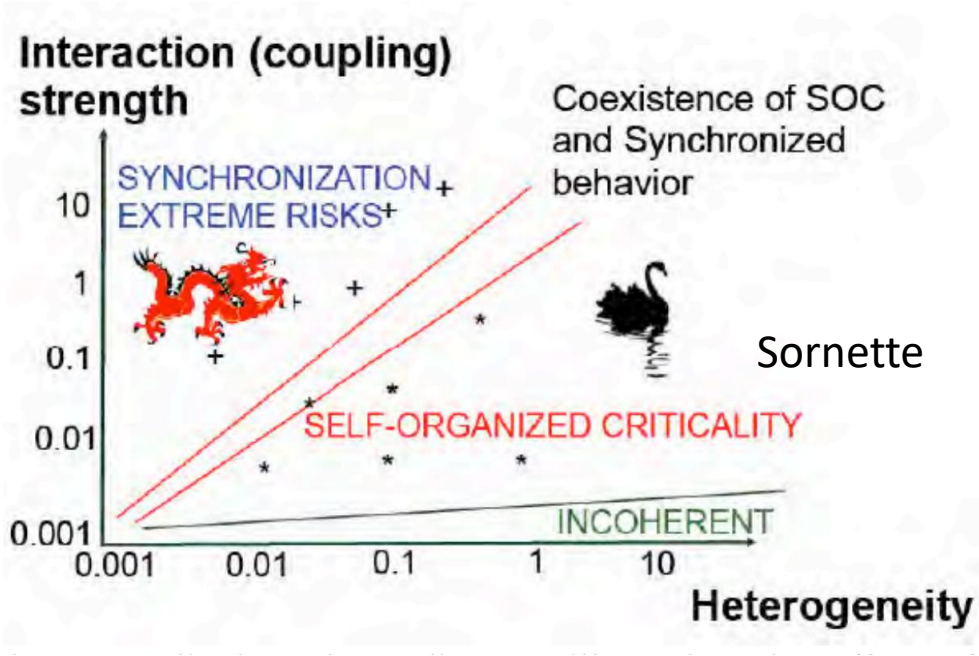
There are many examples in nature and finance where events occur that are outside the range of expectation

Rogue waves



Amplitude of wave more than twice the upper third of the background





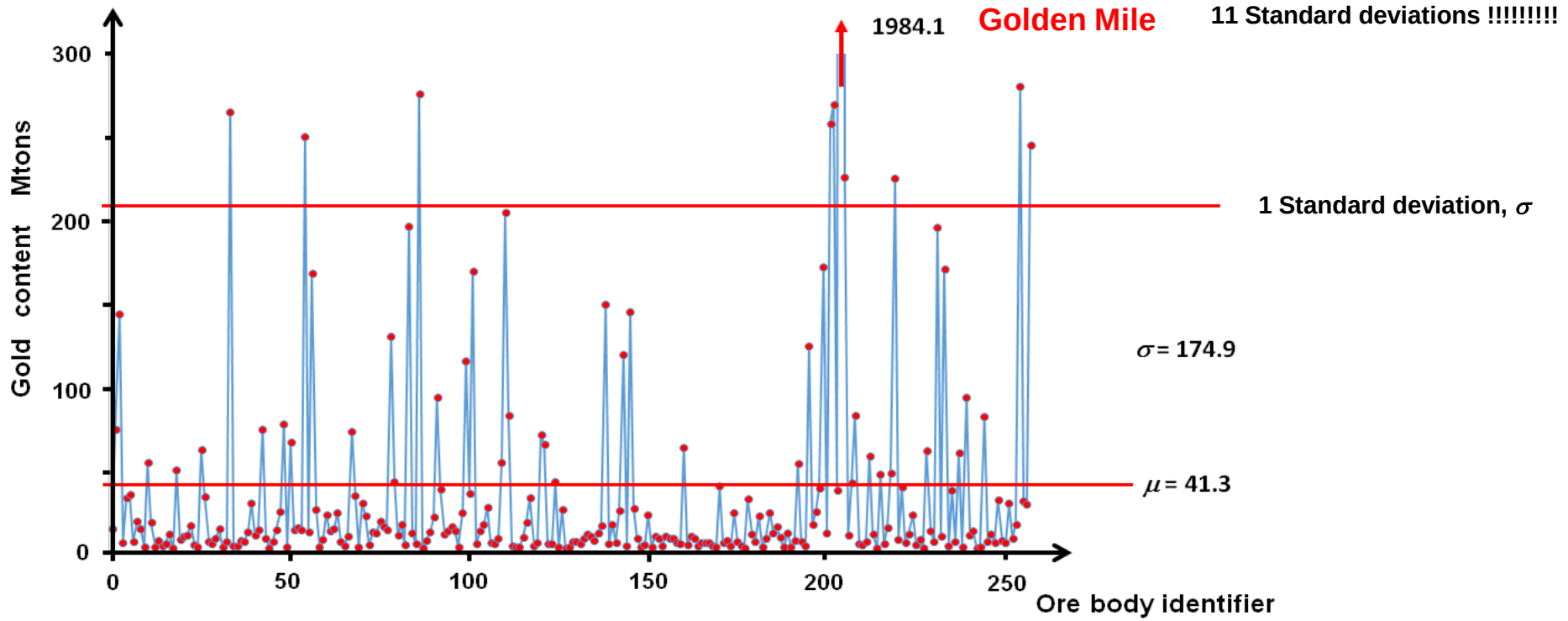
Sornette

Sornette Beyond Black Swans:
The Dragon Kings of Climate Change

The question is: Are giant ore bodies the equivalent of rogue waves or dragon kings?

And if so, does that tell us something about why they are so big?

Distribution of gold endowment in the Yilgarn

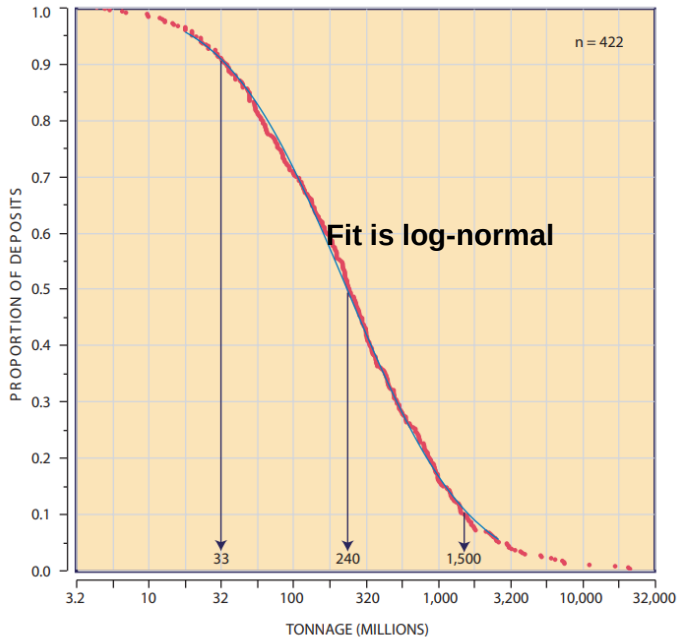


The Golden Mile is way outside expectations and is in the realm of dragon kings and rogue waves.

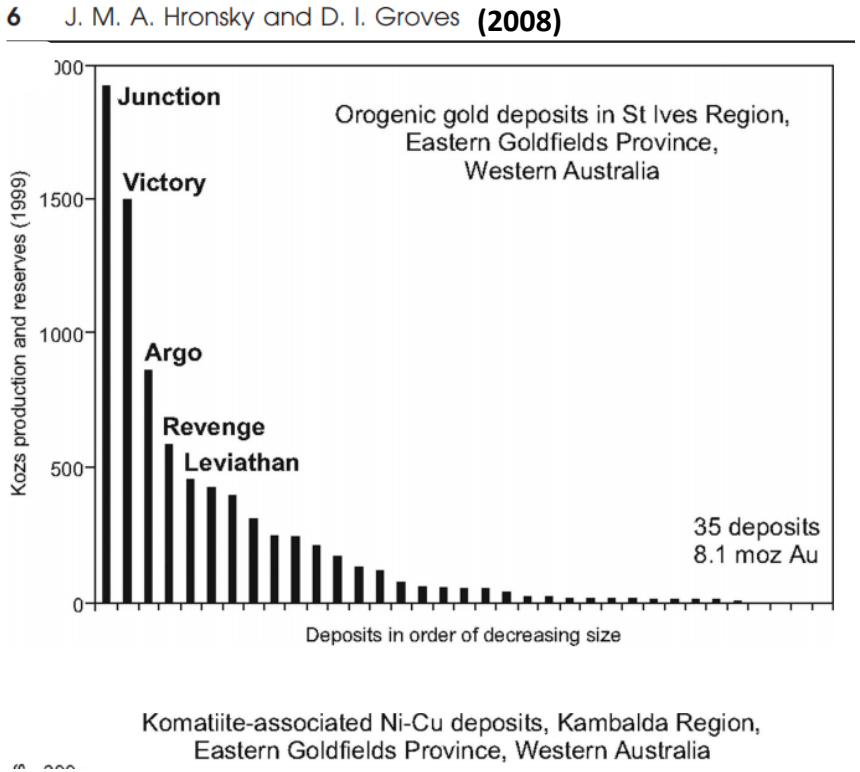
There are two current ways of thinking about mineral endowment:

One is to propose that the endowment follows a log-normal distribution.

The other is to propose a power law distribution (sometimes expressed as a Zipf distribution)



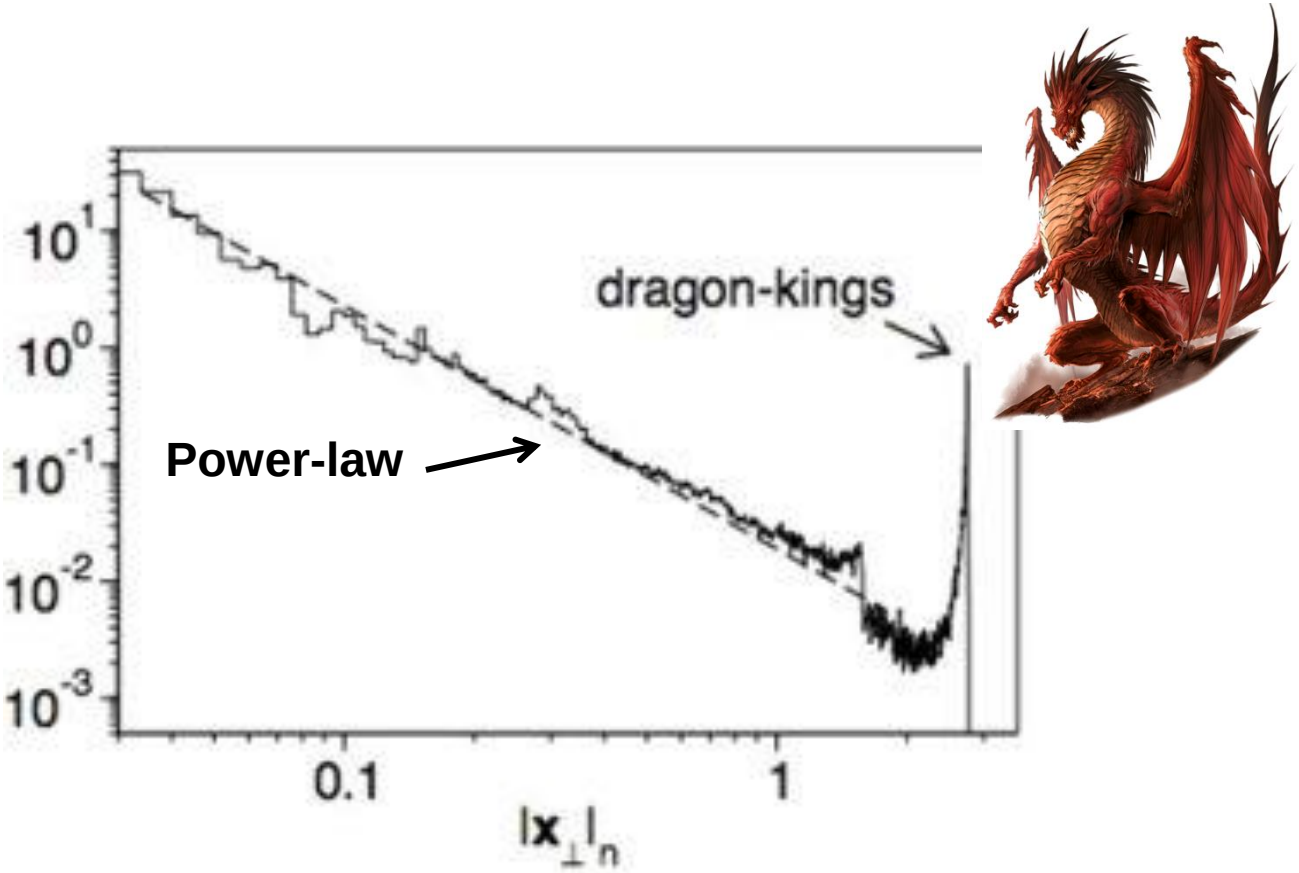
Porphyry copper
From Singer, Berger and Moring, 2008.



There is little attempt to rationalise the difference between the two types of distributions or to show that the proposed distributions are best fit or even true.

We note that dragon kings are not parts of power law distributions.

So it seems that giant ore bodies belong to statistical distributions that are different to power laws (or log-normal: “thin” fat tail).



Fundamental exploration questions:

1. **How do I make a decision using limited data? Even if the data consists of alteration assemblages rather than mineralisation?**

2. **What is the most efficient data collection strategy? What is the minimum amount of data needed?**

3. **What is the most appropriate risk curve? When to quit? Beta distribution**

4. **How do I quantify uncertainty? Entropy analysis.**

5. **What is noise? What is an anomaly? What is and is not an outlier? What is a false positive?**

6. **How do I quantify the difference between mine site, brown fields and green fields exploration?**

7. **How do I quantify prediction? How do I quantify patterns?**

8. **What is the probability of discovering a world class ore body in this region?**

The five fundamental exploration tenets:

1. **Mineralising systems should be viewed as living vibrant open entities far from equilibrium. They are not dead closed linear systems at equilibrium Hence we need *tools, languages and concepts* that take this into account.**

2. **Discovery is an extreme event. A discovery event is many standard deviations from the norm. Hence *extreme event statistics* are fundamental.**

3. **If you cannot measure it you cannot manage it or communicate it. Precise and rigorous science based language is fundamental. This also means that scientifically and information based *measurement standards* must be enforced.**

4. **New paradigms do not result from consensus. Hence leading edge companies must be prepared to be *scientific risk takers*.**

5. **Adoption of new paradigms does result from *consensus*.**