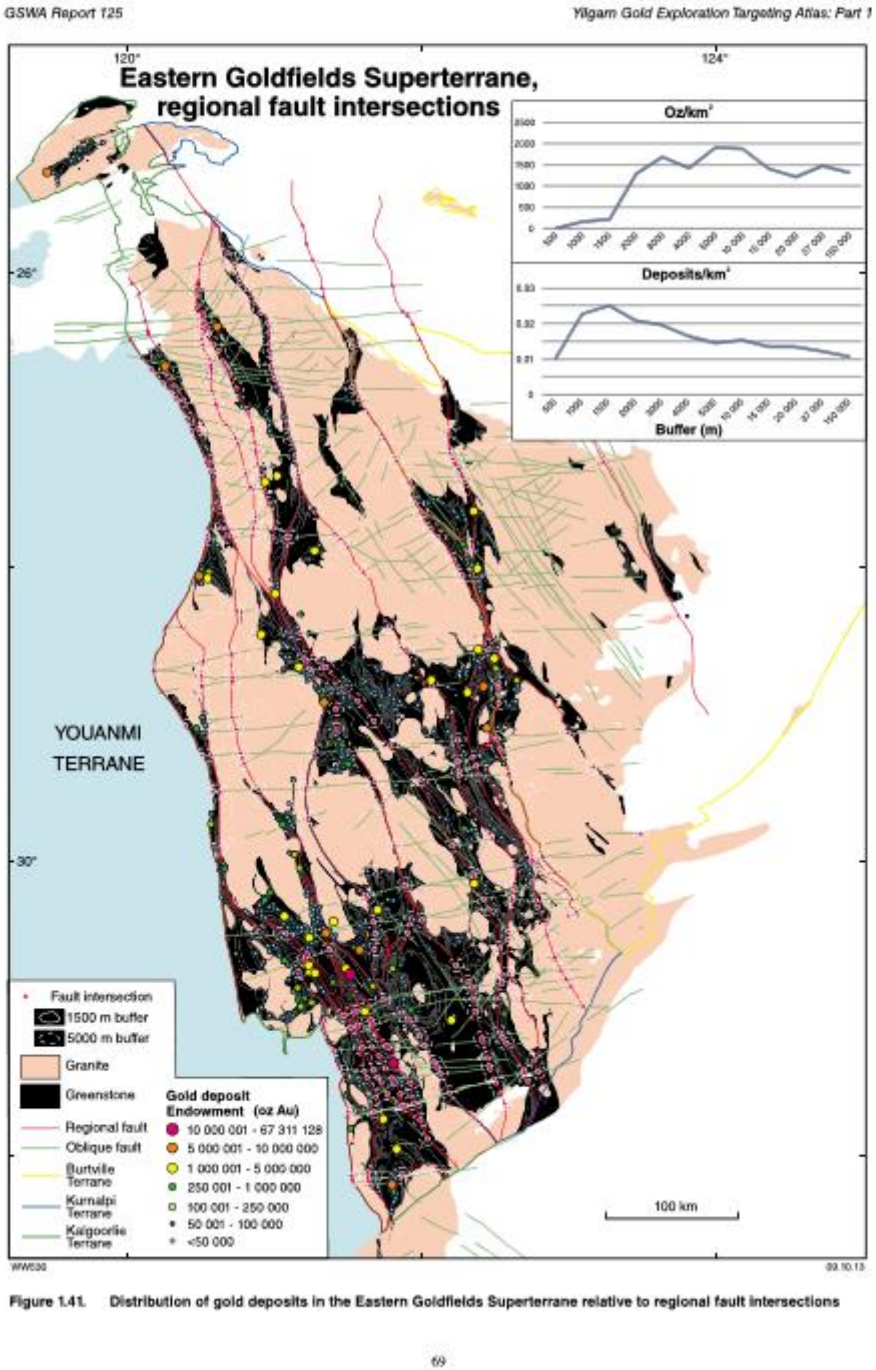
New Tools for Mineral Systems:

Data driven non-parametric analytics

Bruce Hobbs[1,2] and Alison Ord[1]
[1]Centre for Exploration Targeting, UWA.
[2]CSIRO. Perth

Invited keynote for Session 3.16 AGCC: New frontiers in ore systems research.



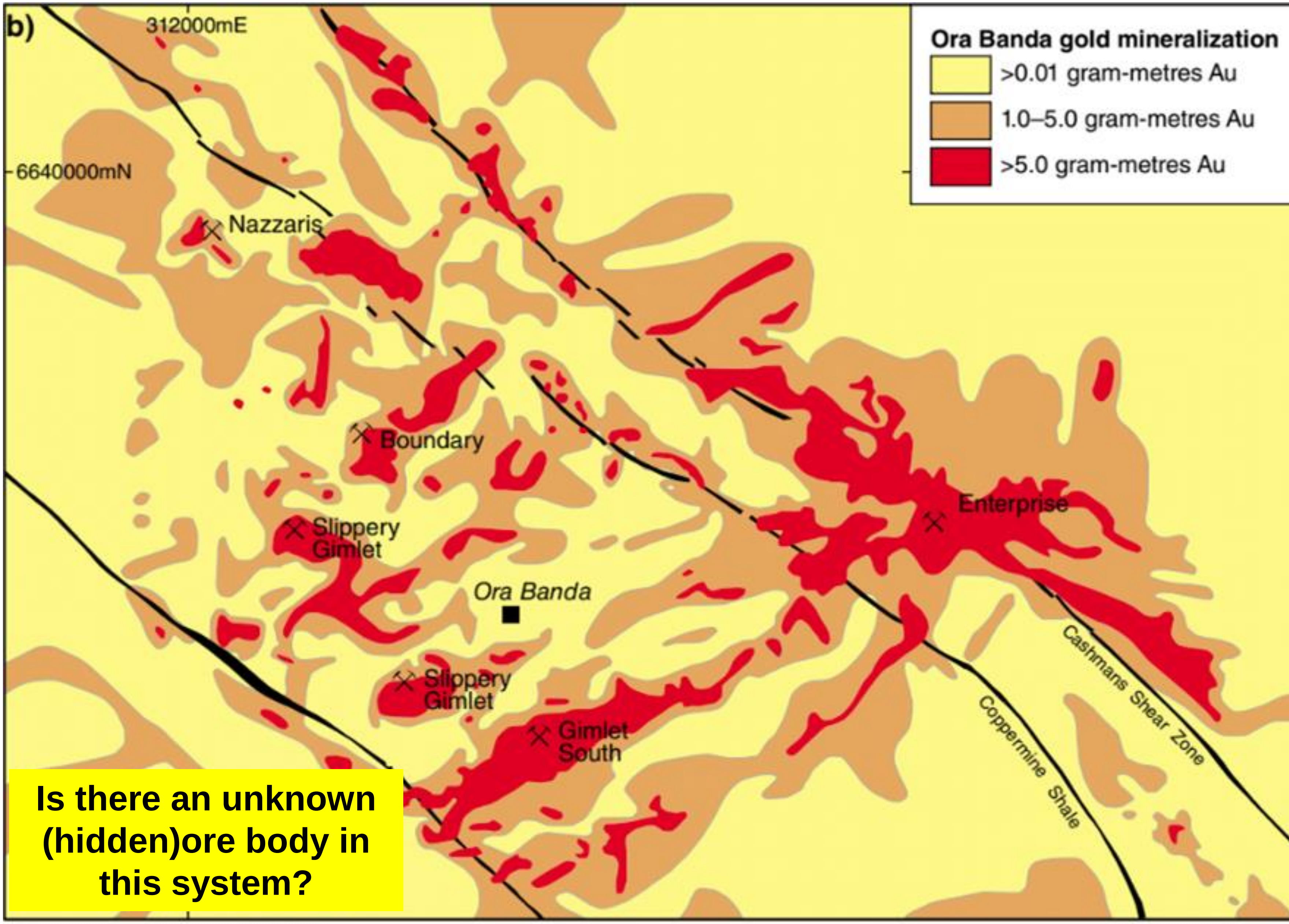
Witt et al., 2014.

Many methods used to create “prospectivity maps”

Weights of evidence, kriging, linear and nonlinear regression

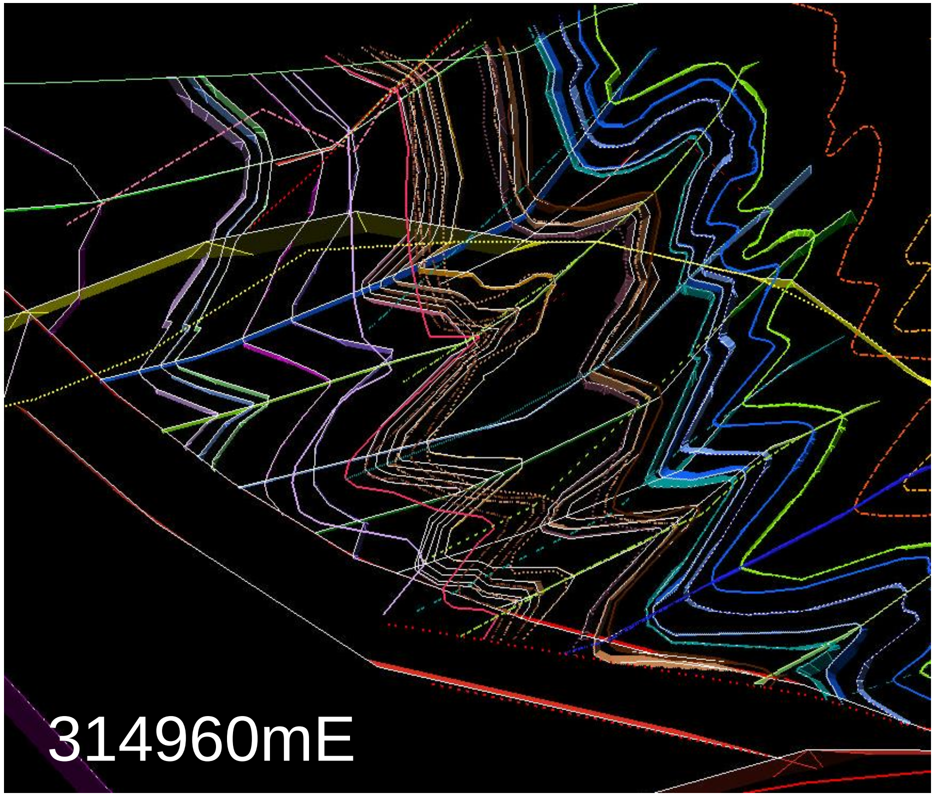
All are linear parametric methods

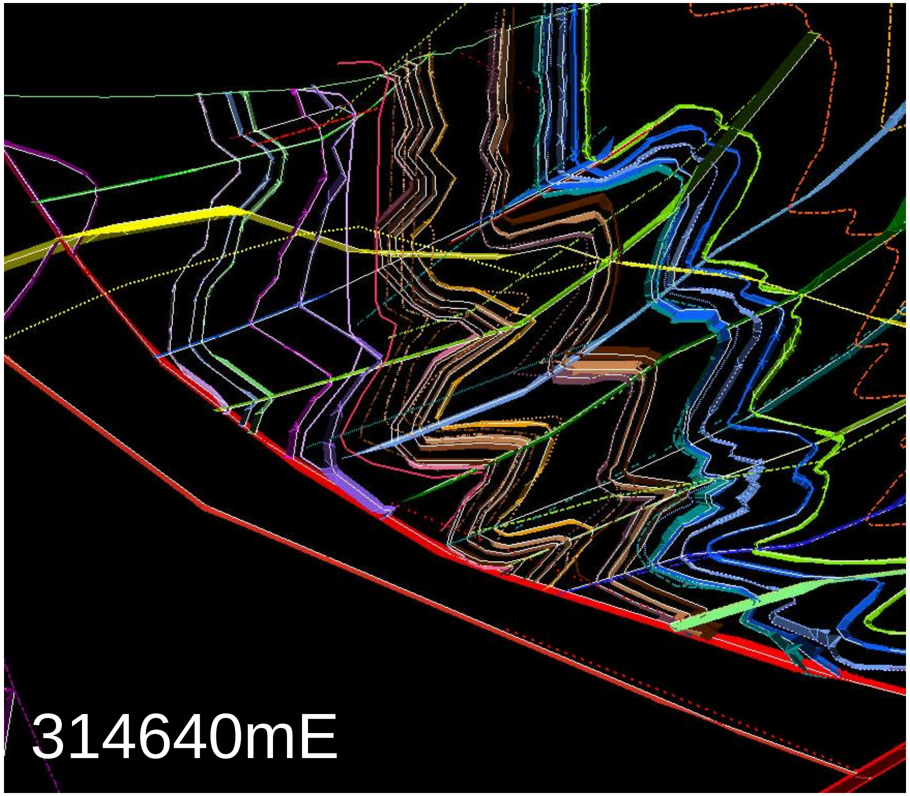
They have no relation to the processes that produced the mineralisation



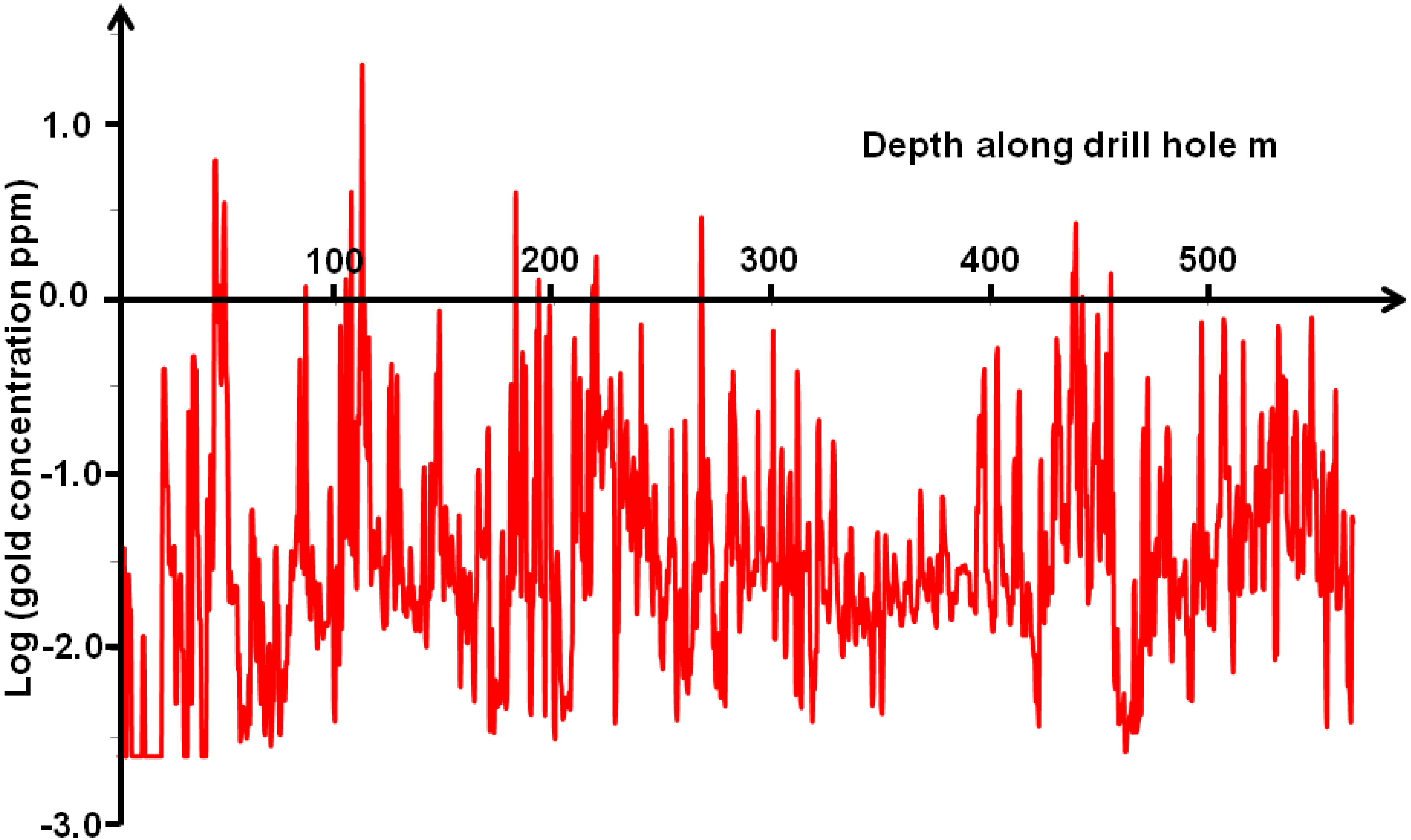
From Witt et al., 2016. After Tripp and Vearncombe, 2004.

Prediction in mining geology





320 m between drilled sections. How do we predict between sections? What is the optimum spacing of drill holes to enable prediction?



What is the grade over the next 10m?

It would be handy to know if there was some systematic way of thinking about things in nature that are intrinsically unforecastable and of ones where reasonable forecasts can be made.

And if things are forecastable are there ways of quantifying the uncertainty?

That is what this talk is all about.

Let us define some terms:

Parametric: a statistical model (Gaussian, log-normal) is assumed for the data.

Nonparametric: No model is assumed; the data speak for themselves ("data driven").

Linear system: Small inputs result in small outputs.
Fourier transforms useful.

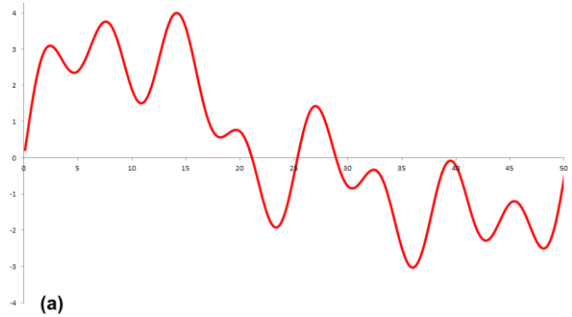
Nonlinear system: Small changes result in unexpected or large results.
Fourier transforms not useful.

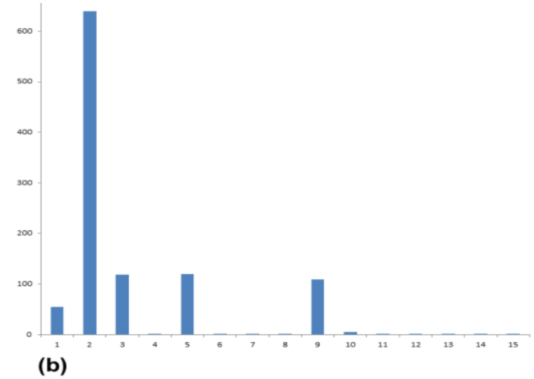
Stochastic: The result of uncorrelated processes (tossing a coin).

Deterministic: The result of well defined laws.

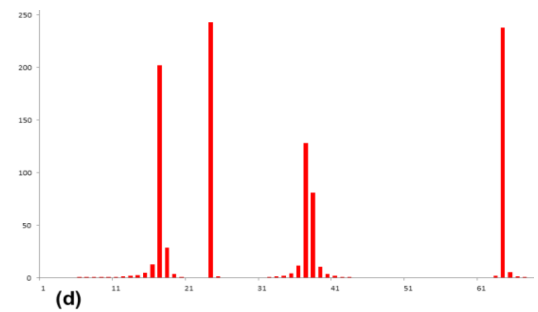
Signal | **Fourier transform**

Periodic



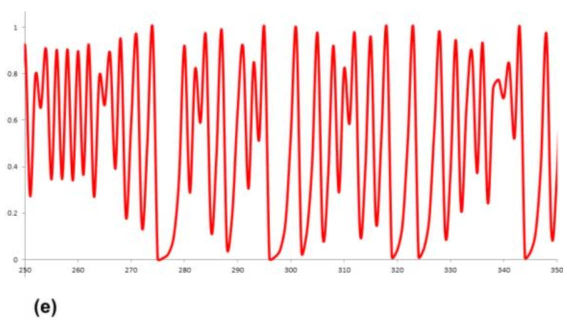


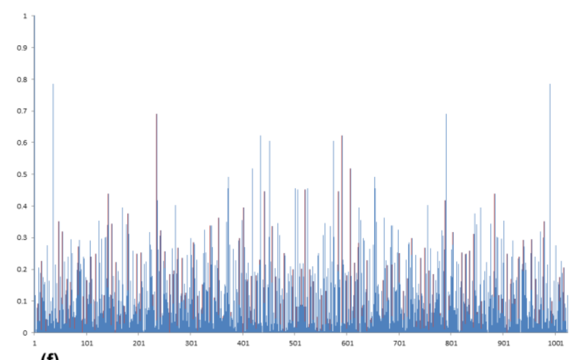
Quasiperiodic



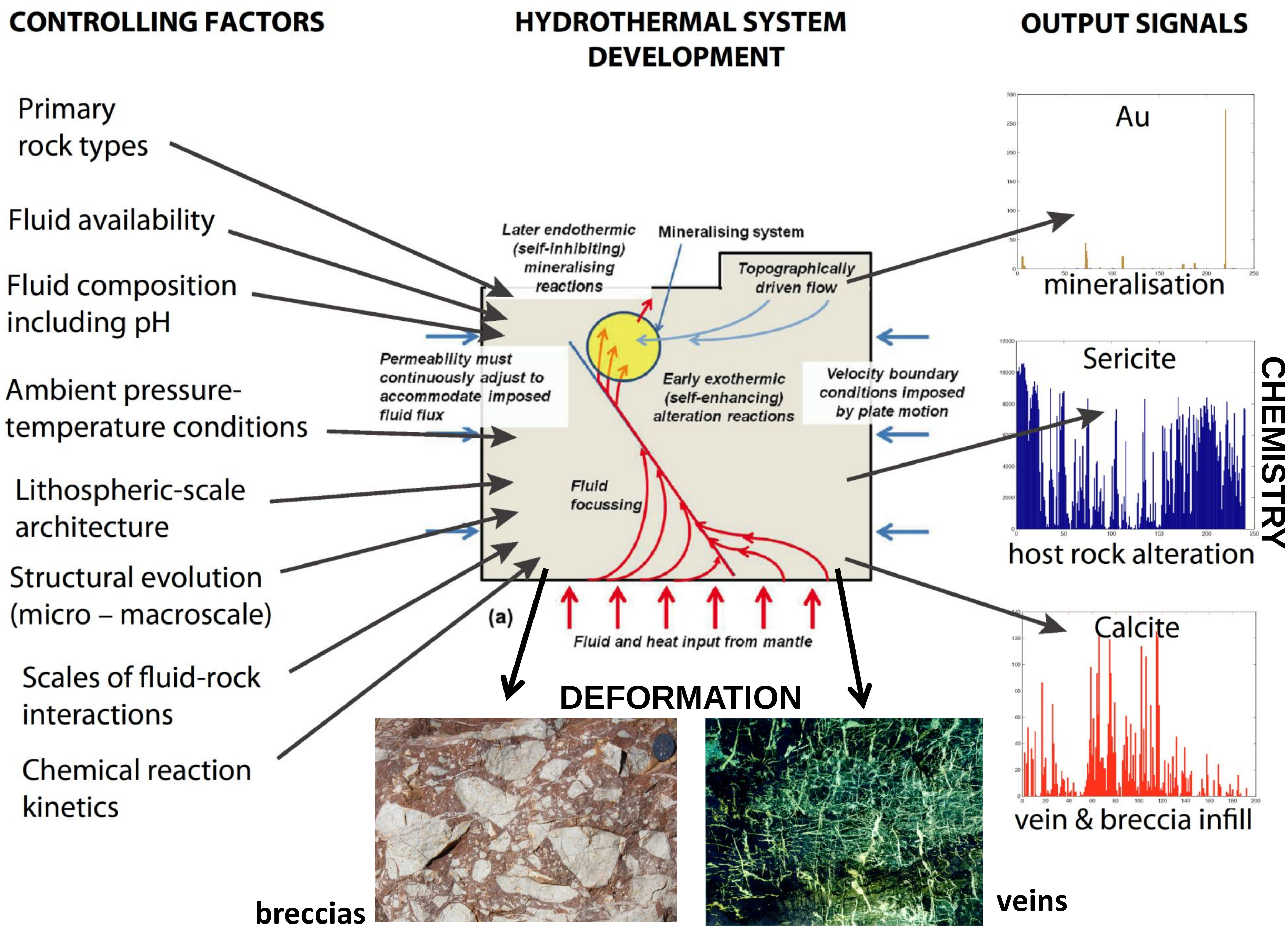
$$y = \sin(x) + \sin(\sqrt{2}x) + \sin(\sqrt{5}x) + \sin(\sqrt{15}x)$$

Chaotic





$$x_{n+1} = 4x_n\left(1 - x_n\right)$$



CONTROLLING FACTORS
HYDROTHERMAL SYSTEM DEVELOPMENT
OUTPUT SIGNALS
Primary rock types
Fluid availability
Fluid composition including pH
Ambient pressure-temperature conditions
Lithospheric-scale architecture
Structural evolution (micro – macroscale)
Scales of fluid-rock interactions
Chemical reaction kinetics
Later endothermic (self-inhibiting) mineralising reactions
Mineralising system
Topographically driven flow
Permeability must continuously adjust to accommodate imposed fluid flux
Early exothermic (self-enhancing) alteration reactions
Velocity boundary conditions imposed by plate motion
Fluid focussing
(a)
Fluid and heat input from mantle
Au
mineralisation
Sericite
host rock alteration
CHEMISTRY
Calcite
vein & breccia infill
DEFORMATION
breccias
veins

This presentation outline some new tools useful in studying mineral systems.

The established tools include the traditional workhorses which should not be underestimated:

The historical development of the subject involves

- The development of geophysical techniques, particularly airborne surveys.
- The development of geochemical techniques.
- Improvements in drilling technologies.
- The development of mineral mapping technologies at drill hole and regional scales
- The application of statistics to the definition of prospectivity maps and to ore grade estimation and the development of block models.
- The application of equilibrium thermodynamics to understanding the chemical evolution of mineral deposits.

The continued development of these tools has resulted in an avalanche of data.

We now call this agglomeration of data “big data” and various popularised data analysis technologies have been developed in order to extract some meaning from these data sets.

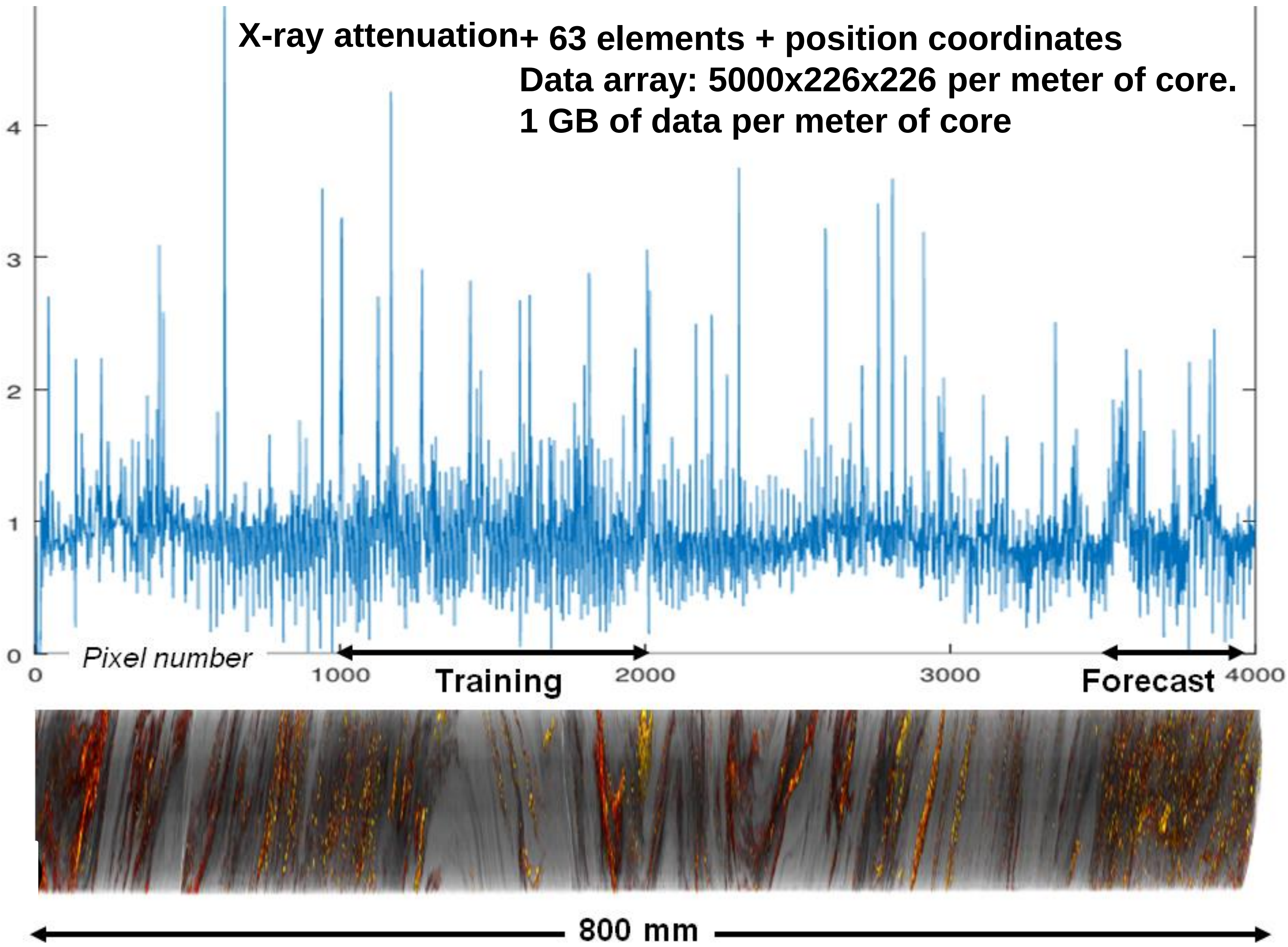
These data analytic procedures go under names such as:

- **Artificial intelligence.**

- **Deep learning.**

- **Unsupervised deep neural networks.**

None of these approaches involve the dynamics of the processes that produced the data in the first place and hence are somewhat *ad hoc*.


X-ray attenuation+ 63 elements + position coordinates
Data array: 5000x226x226 per meter of core.
1 GB of data per meter of core
0
1
2
3
4
Pixel number
0
1000
2000
3000
4000
Training
Forecast
800 mm

Even the simplest of nonlinear systems can generate behaviour of great complexity.

But even very complex, even chaotic, behaviour is not arbitrary. The behaviour is highly constrained by the underlying physical and chemical processes that operated to produce the system.

The traditional way of exploring such systems is the write the differential equations that describe such nonlinear behaviour and then try and understand the behaviour of such equations.

A relatively new, approach is to adopt data driven analytics where the data is used to derive the equations.

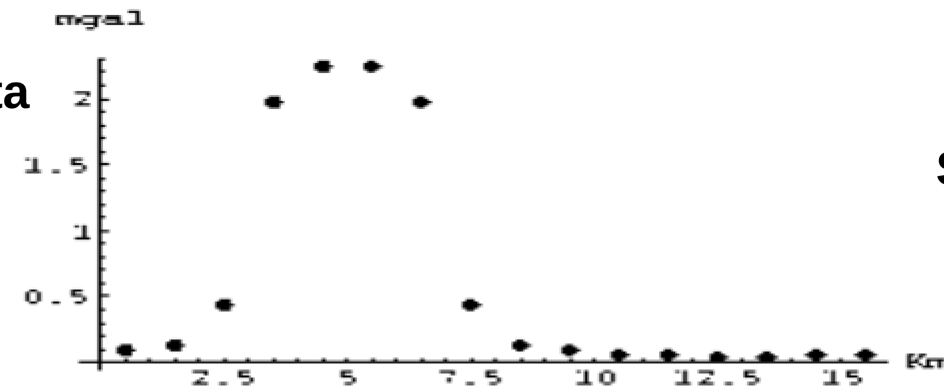
An important advance in data driven approaches is the recognition that many nonlinear systems can be approximated by linear systems (Koopman transforms) with a change in the type of data collected (for example, the gradient of the chemical composition or the gravity tensor rather than the composition or gravity alone).

Ambiguity in the geosciences arises from linear thinking

The production manager asks a geologist, engineer, and geophysicist what 2 + 2 is. The geologist thinks for a bit and then says “somewhere between 3 and 5”. The engineer fiddles with a calculator and says “3.9999999”. The geophysicist looks her in the eye and asks “what answer do you want”

This is not just an accurate expression of the ways in which geophysicists and geologists operate, it is a direct ramification of linear ways of thinking in the geosciences which result necessarily in ambiguity

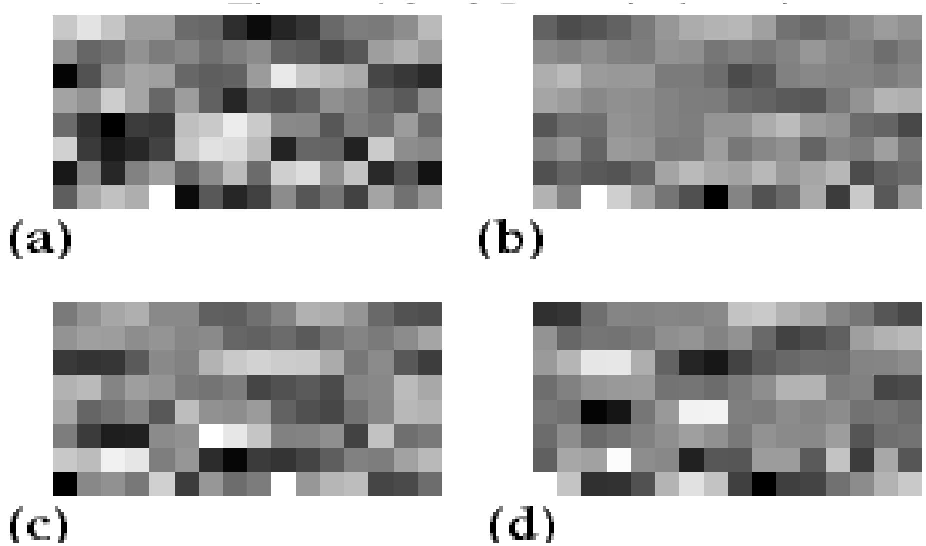
Gravity data



Surface signal

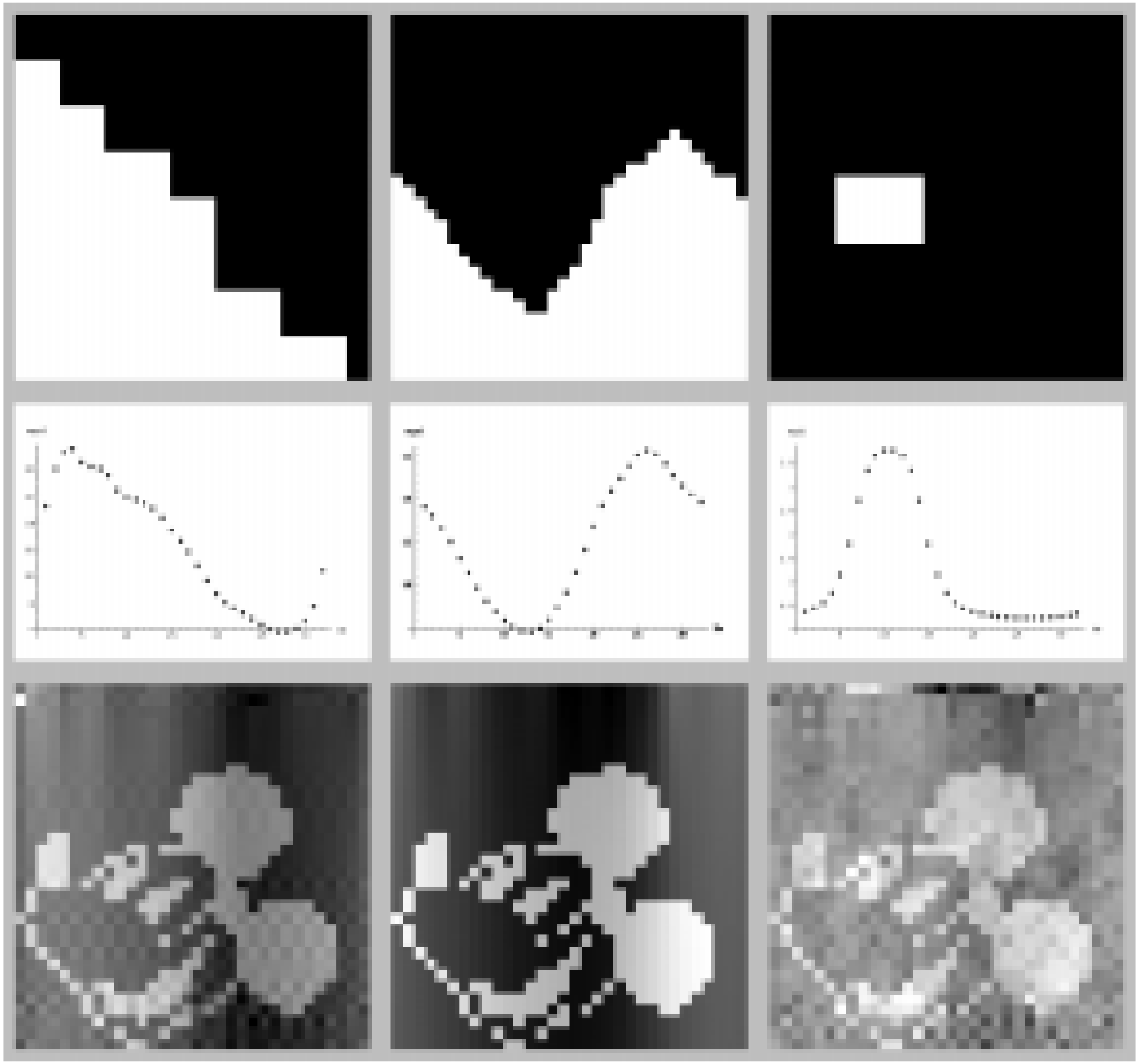
Ambiguity in potential field interpretations

Real situation



Distributions compatible
with observed signal

Mickey mouse fits to potential field data



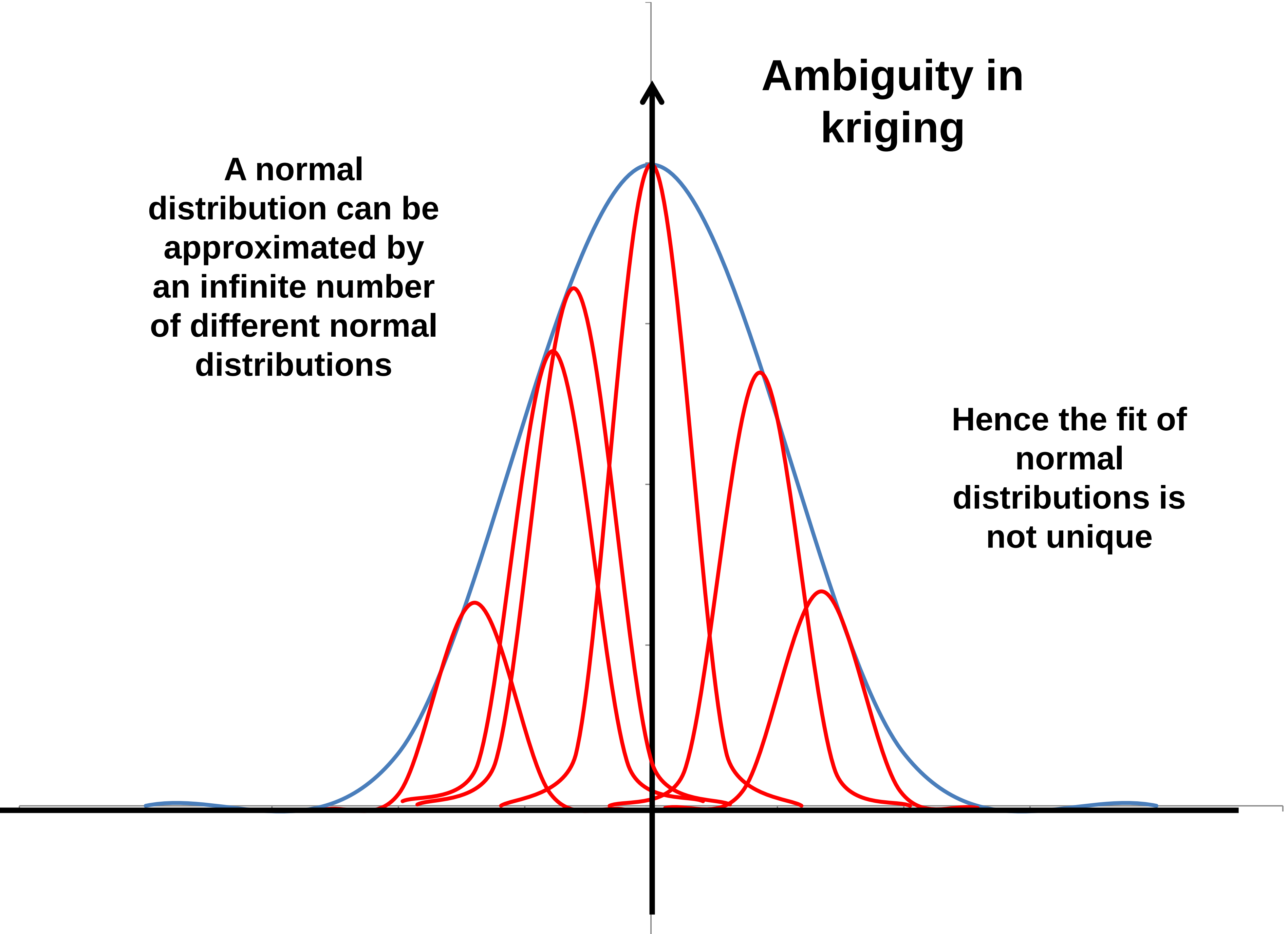
Real situation

Observed signal

Mickey mouse fits to observed signal

Ambiguity in kriging

A normal distribution can be approximated by an infinite number of different normal distributions



Hence the fit of normal distributions is not unique

Ambiguity in simple kriging

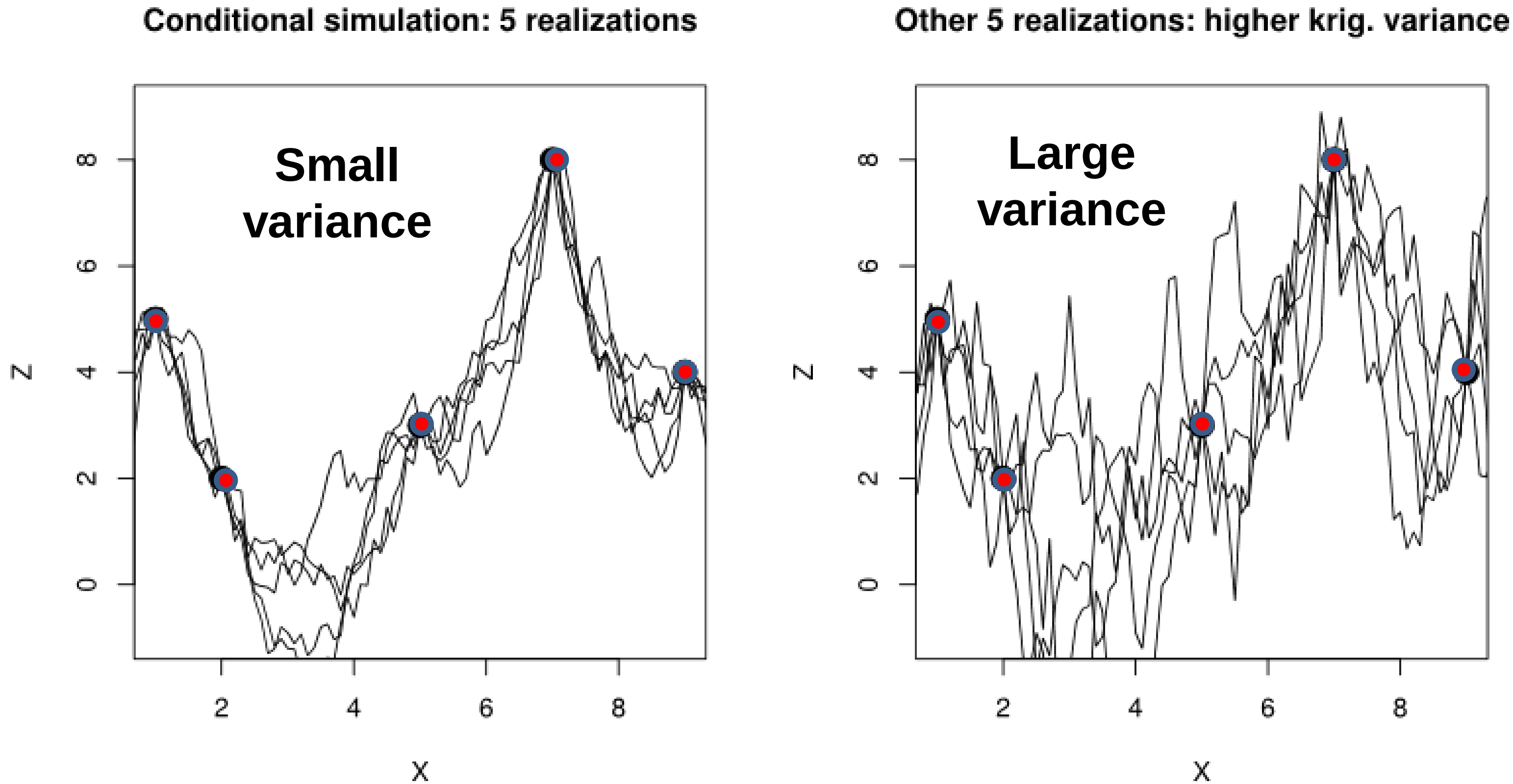


Figure 1: Conditional simulations, data given by dots

All the realisations are faithful to the data and also faithful to the statistical model assumed (mean and variogram).

The governing equations for simple kriging:

$$\mathbf{C}\lambda_0 = \mathbf{C_0}$$

Covariance matrix for all points relative to all points

Covariance matrix for all points relative to a datum point

Since the covariance is ambiguous the solutions to these equations are ambiguous

Ambiguity in seismic interpretations

Seismic wave velocity

$$V_P = \sqrt{\frac{K + \frac{4\mu}{3}}{\rho}} \qquad V_S = \sqrt{\frac{\mu}{\rho}}$$

μ is shear modulus. ρ is density, *K* is bulk modulus

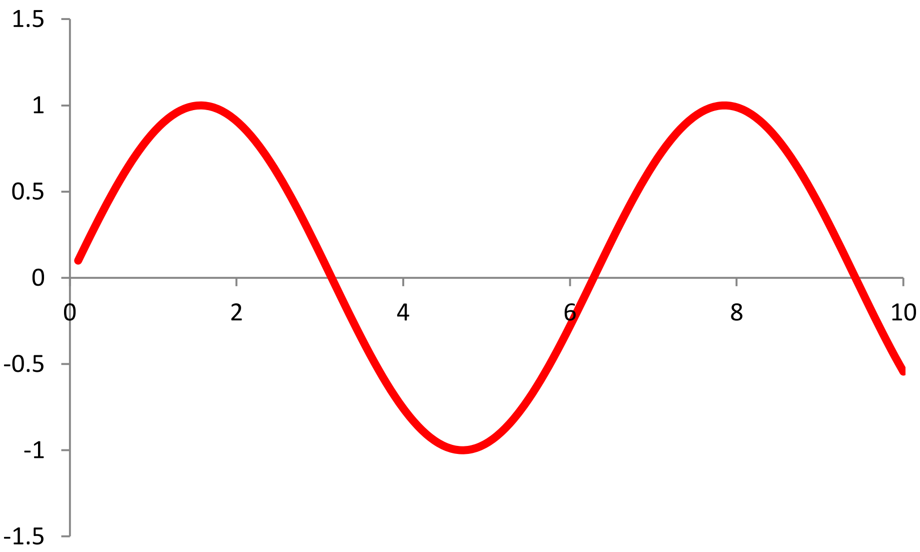
Ambiguity exists since variations in velocity can arise from variations in shear modulus or variations in density or both

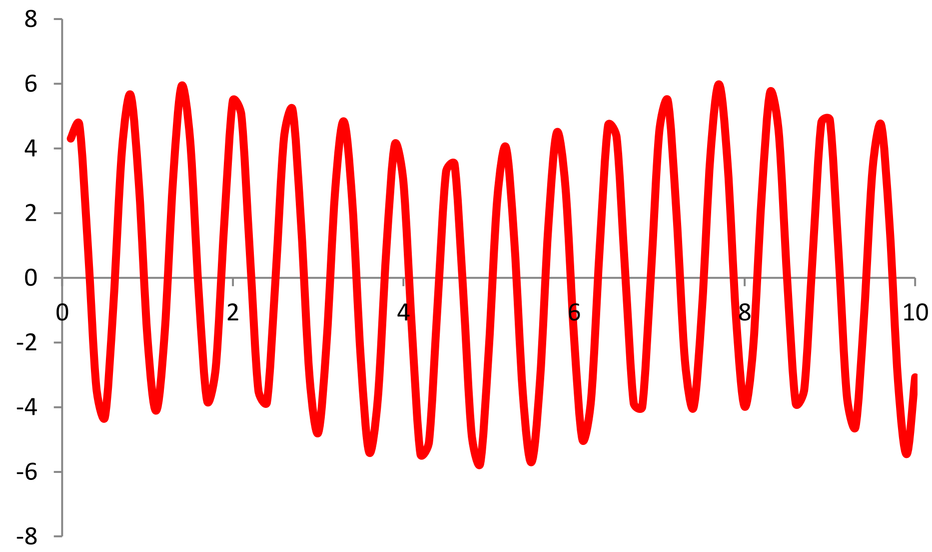
Ambiguity in solution to linear differential equations

If $y = \sin x$

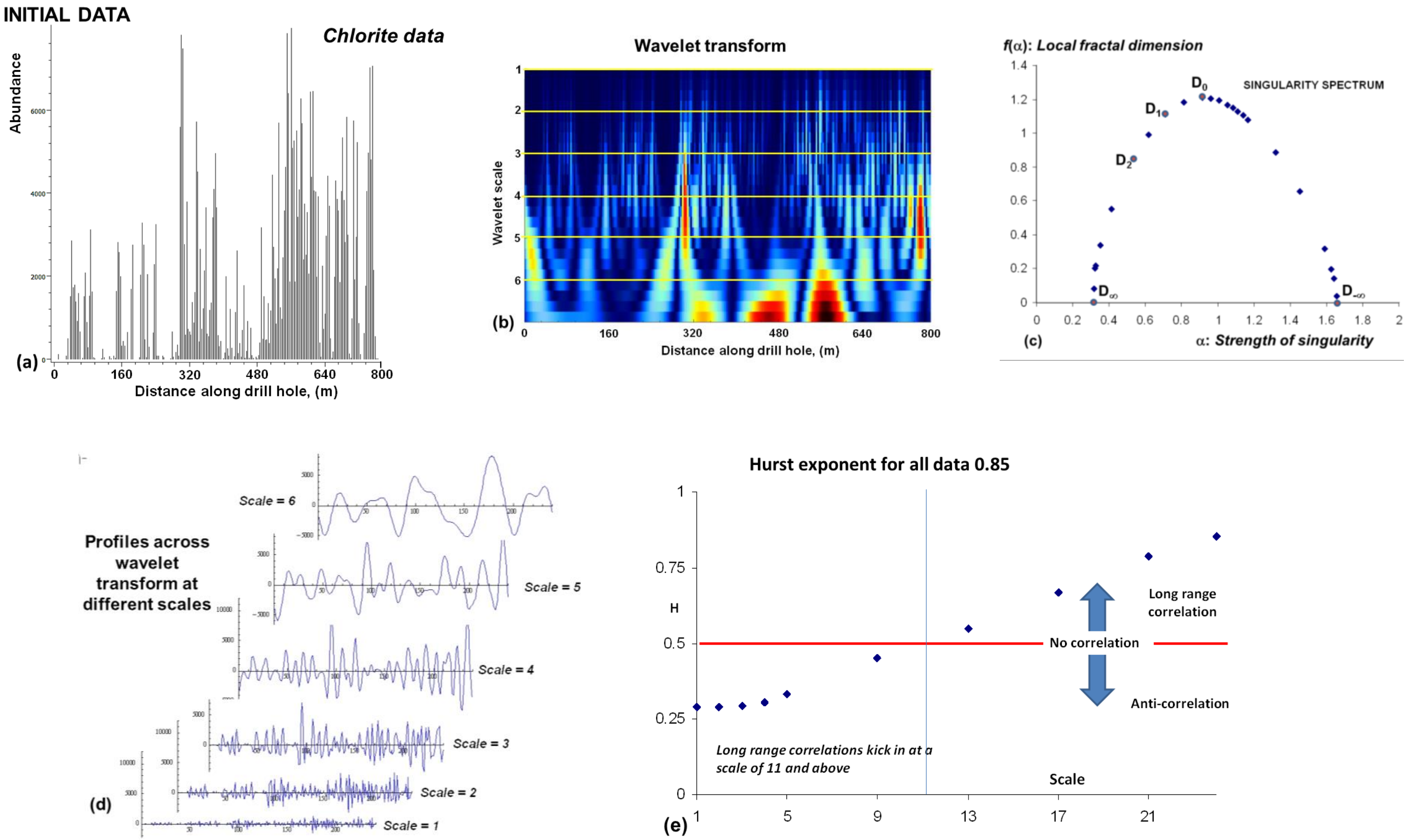
Is the solution to a linear equation, then so is

$$y = \sin x + 5 \sin 10x + \textit{any linear combination}$$





The issue is that most (if not all) geoscience data sets are multifractal.



The issue is that most (if not all) geoscience data sets are multifractal (consist of arrays of singularities).

This includes
- **Mineral distributions**
- **Structures such as folds, joints, veins, breccias and faults**
- **Magnetic susceptibility**
- **Density.**
- **Electrical conductivity**
- **Elastic moduli (shear modulus and bulk modulus)**

Hence any smooth function (Covariance, potential field functions) cannot represent the data which consists of singularities.

The need is for approaches that honour the data.

Takens' theorem (1981):

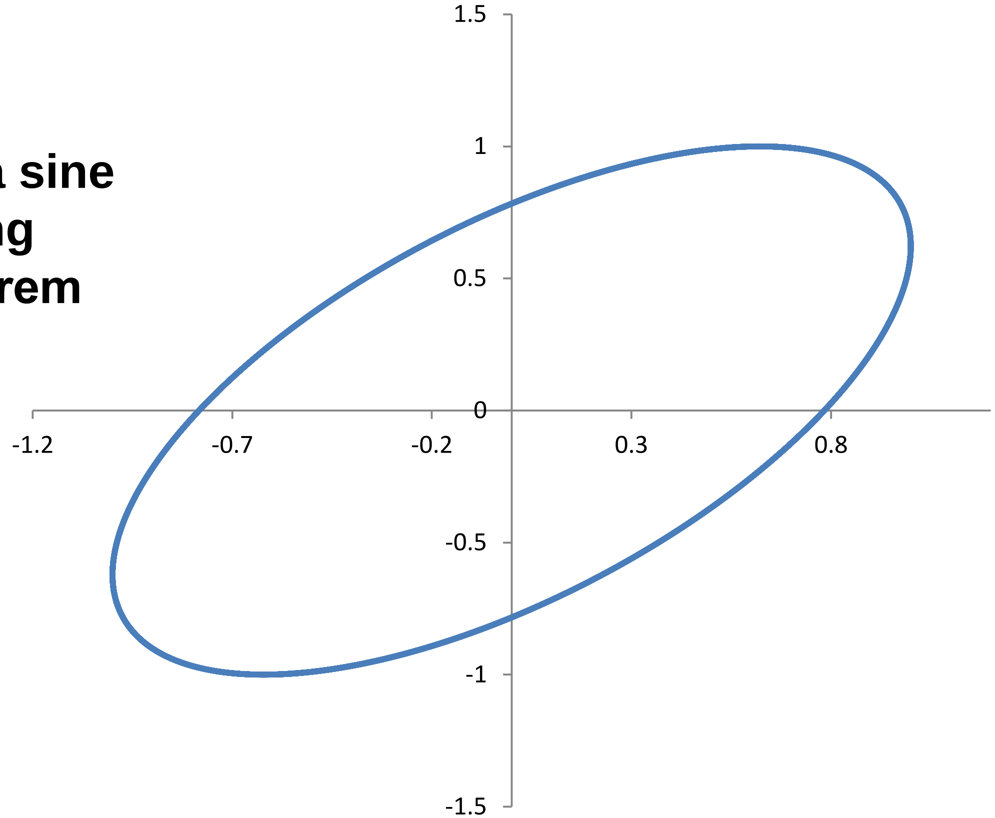
In a coupled nonlinear system all of the behaviour of the system is contained in the behaviour of one component of that system. A component might be the abundance of sericite for instance.

Hence the complete dynamical behaviour of that system can be obtained from one set of data from that system.

The dynamical behaviour is expressed as an *attractor* for that system. and can be derived from one set of data

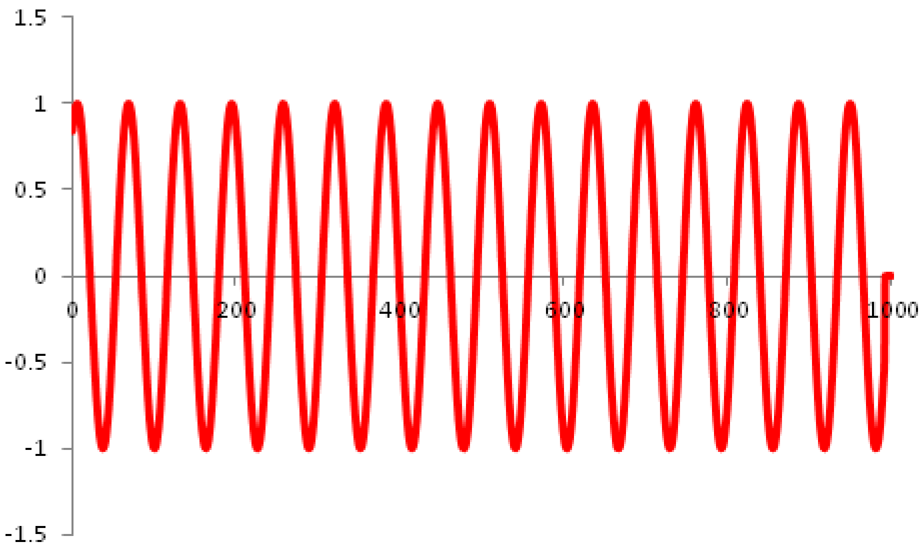
The attractor represents all possible states that the system can occupy.

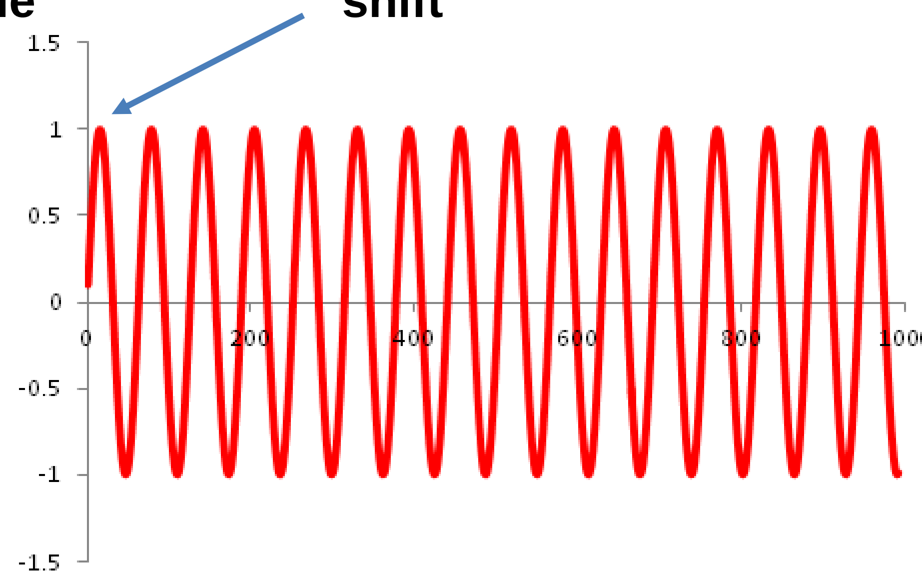
Attractor for a sine curve using Takens'theorem



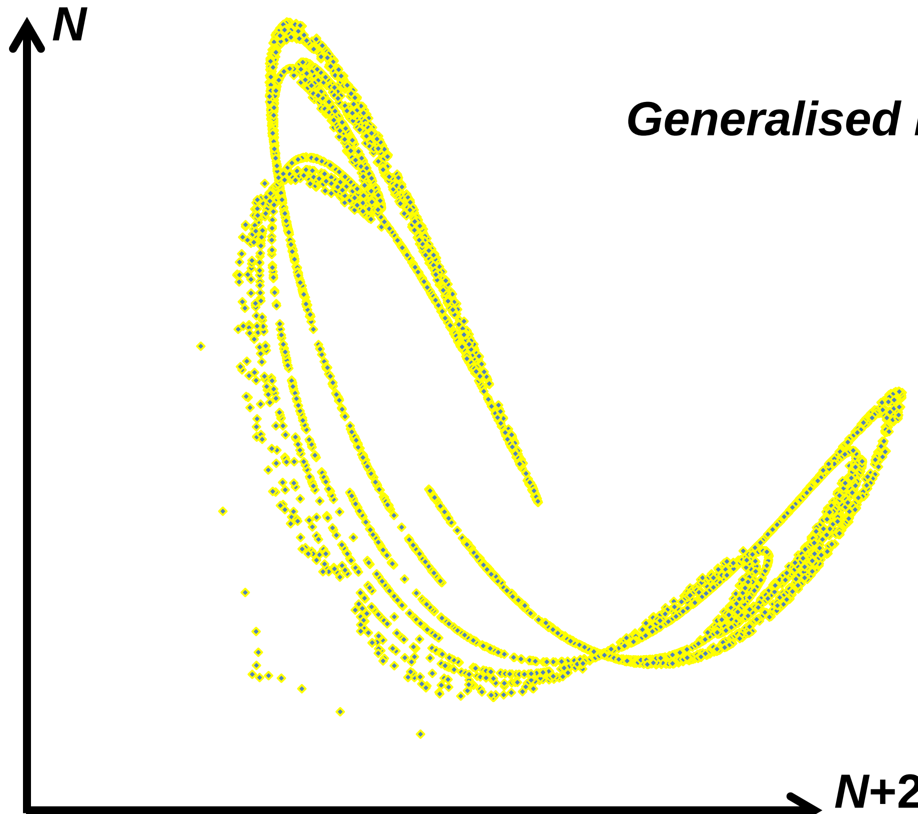
Take the data and shift it a little

shift

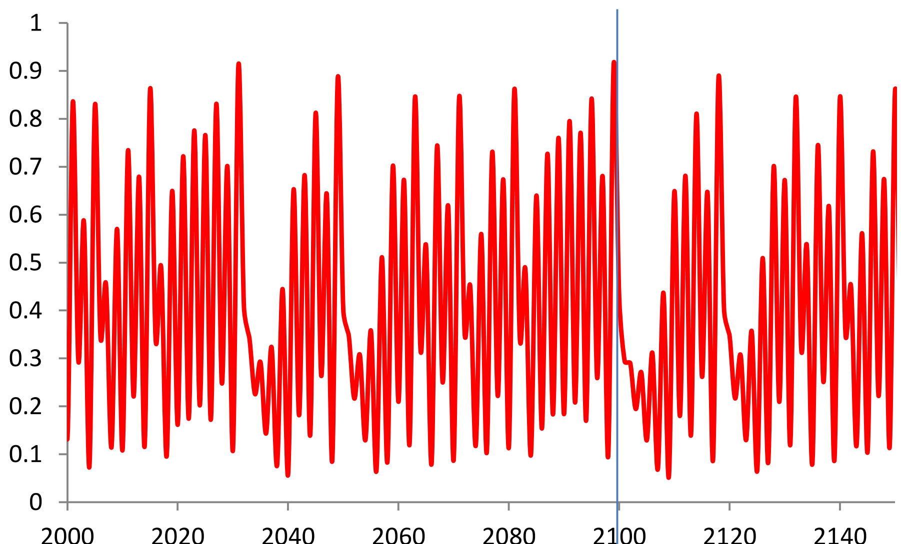


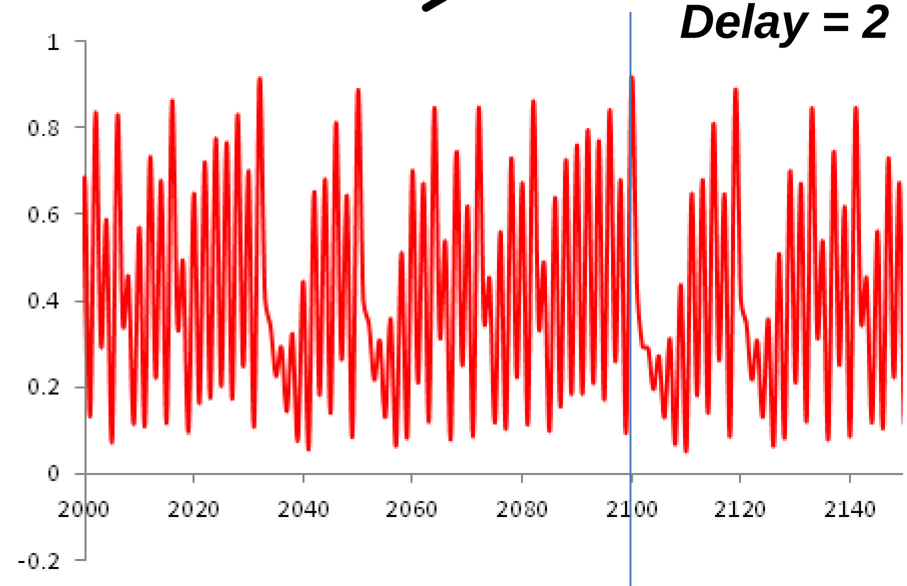


Attractor construction for a chaotic system using Takens' delay theorem

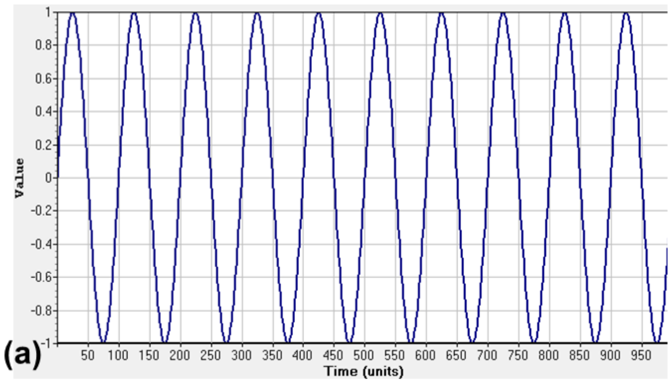


Delay = 2

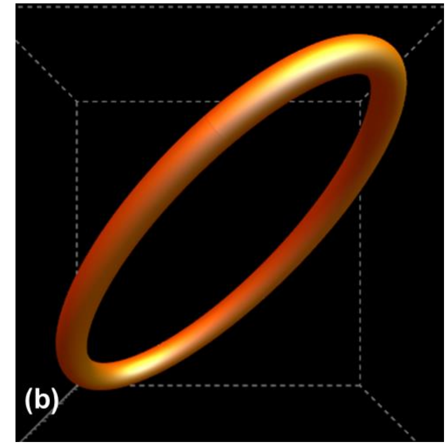


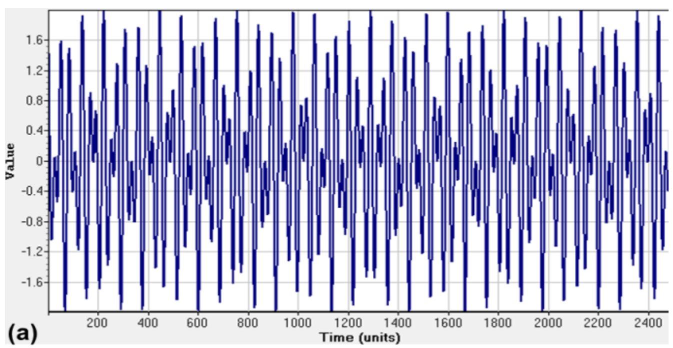


Attractors for various types of systems

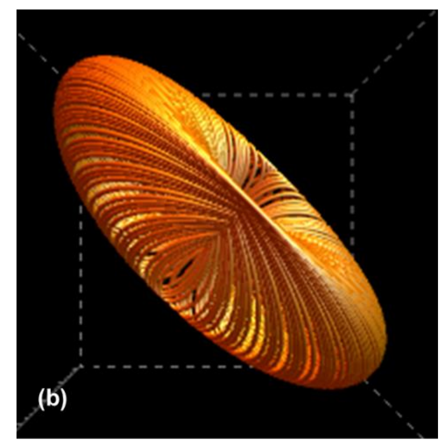


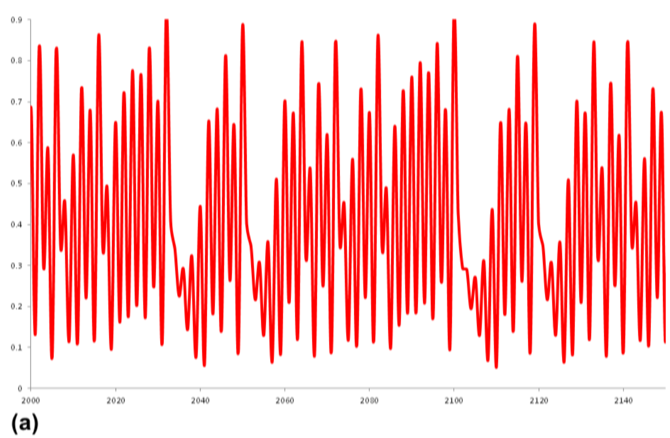
Periodic



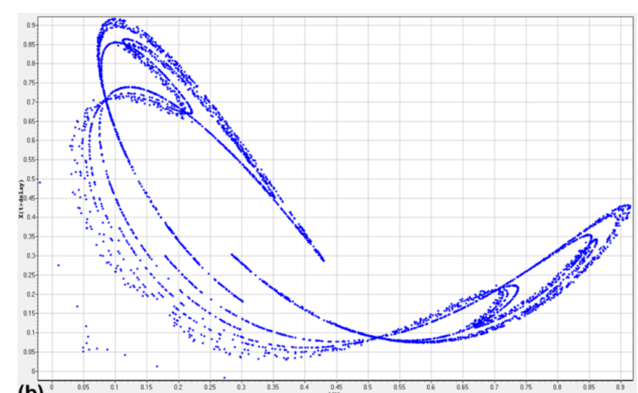


Quasiperiodic



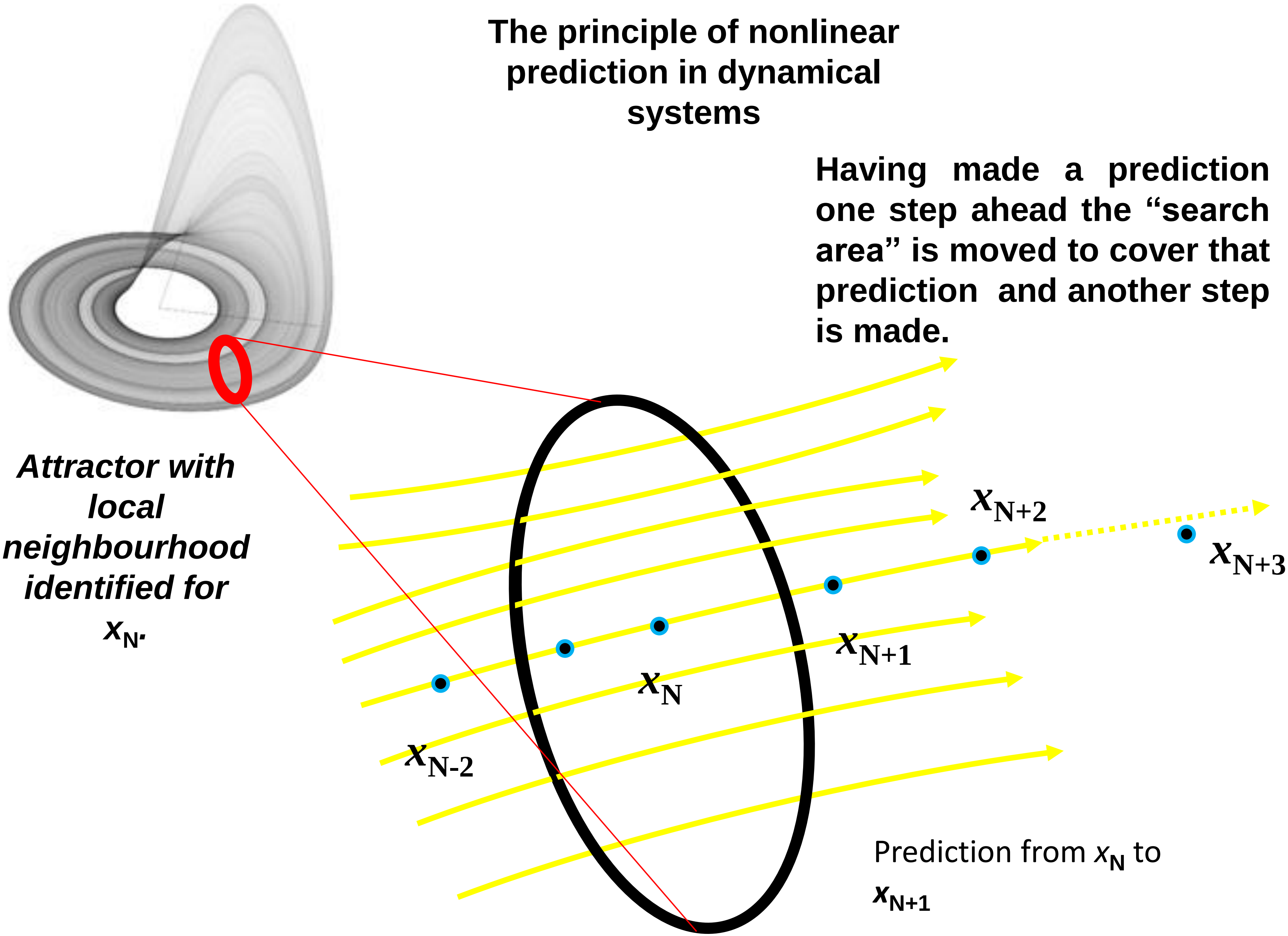


Chaotic

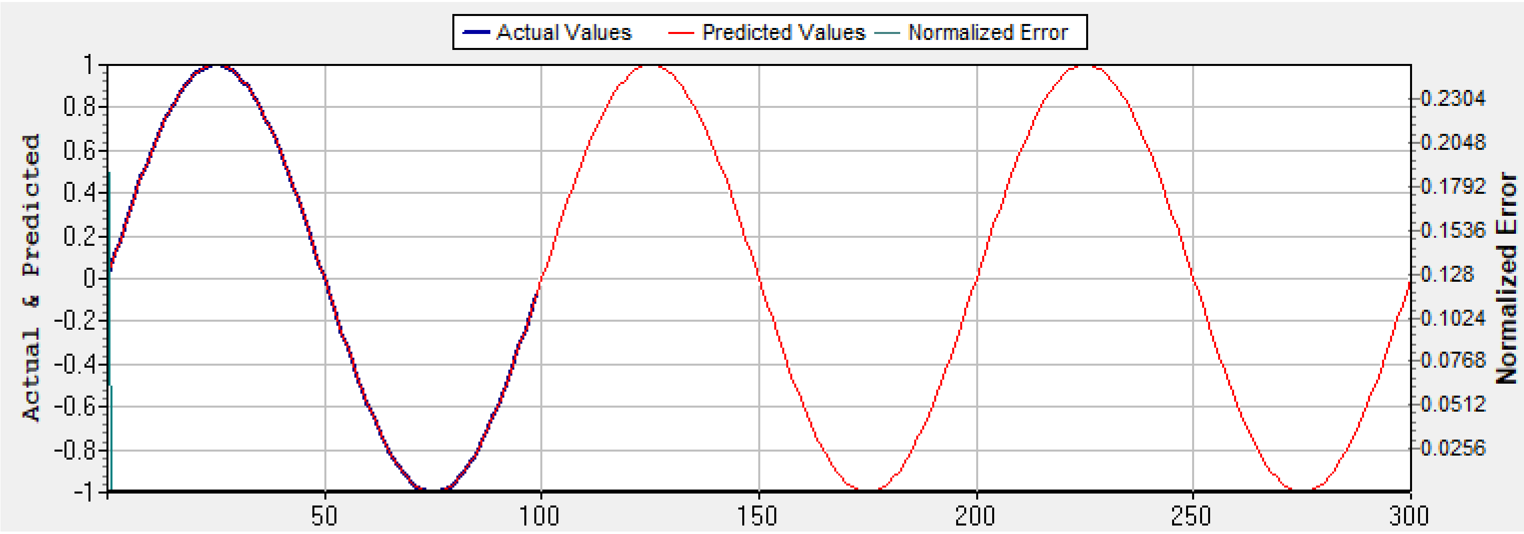


The principle of nonlinear prediction in dynamical systems

Having made a prediction one step ahead the "search area" is moved to cover that prediction and another step is made.



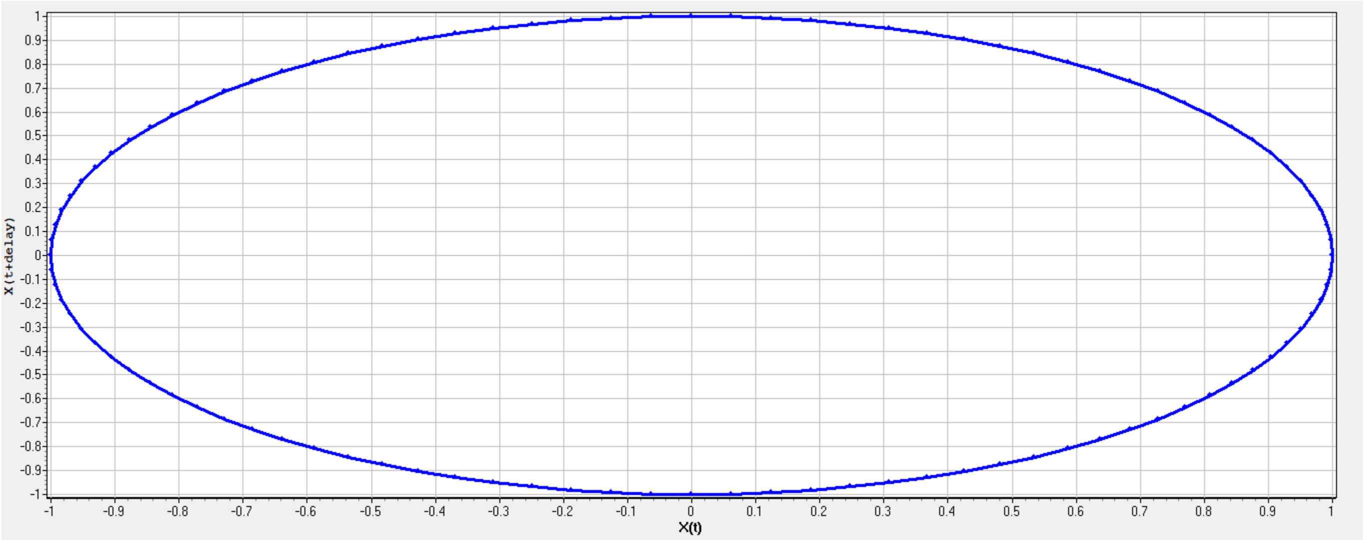
Sine wave with no noise



Train

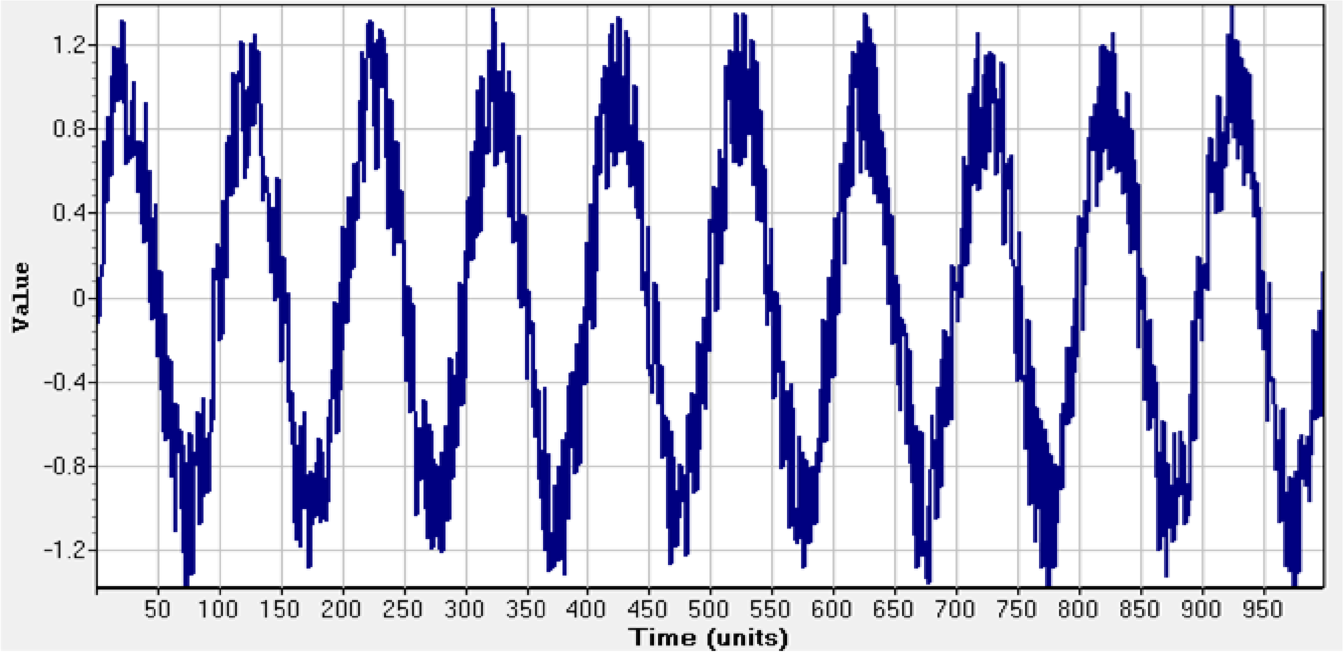
Test

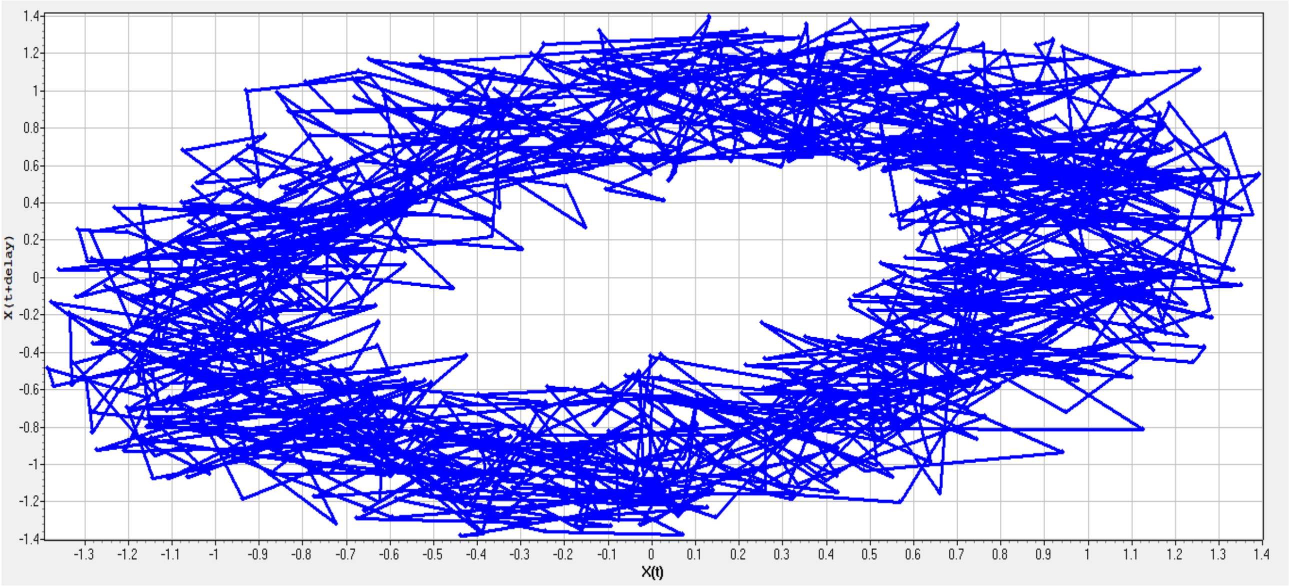
Predict



Attractor

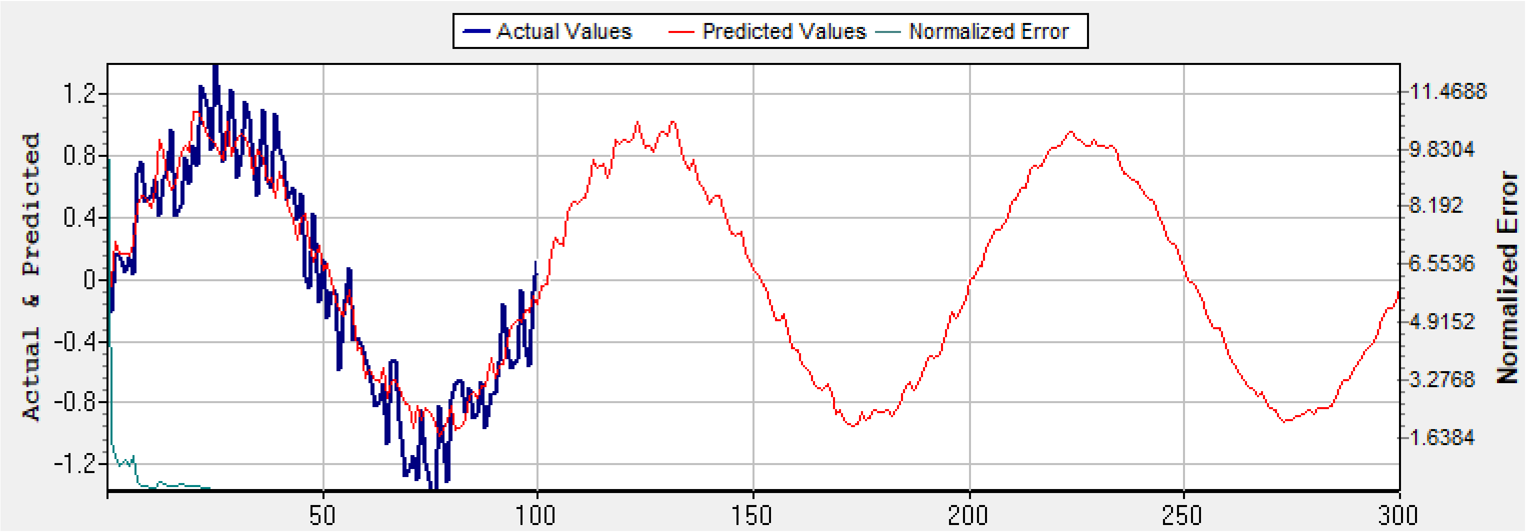
Sine wave with noise





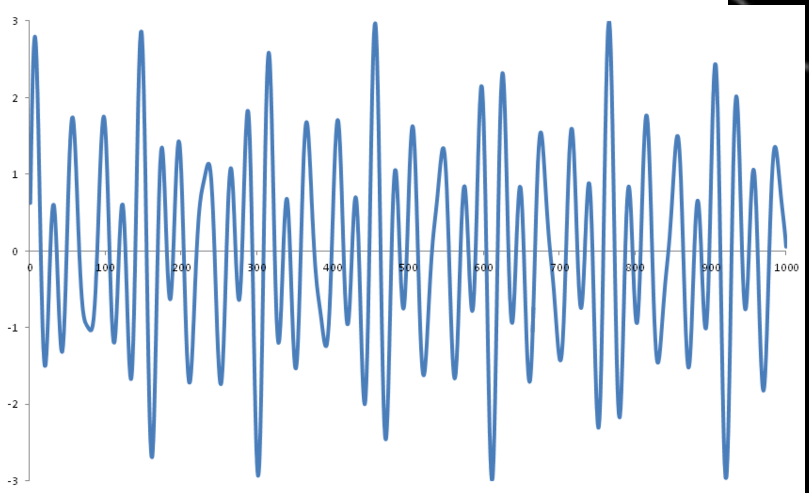
Attractor

Sine wave with noise

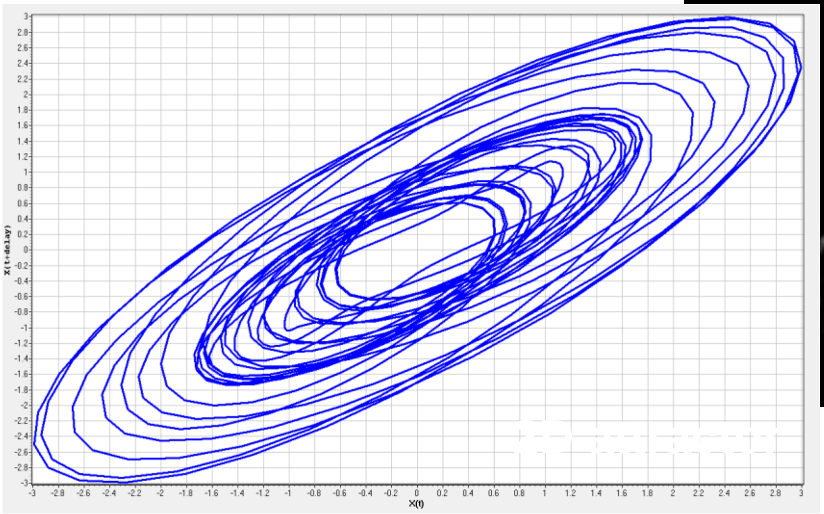


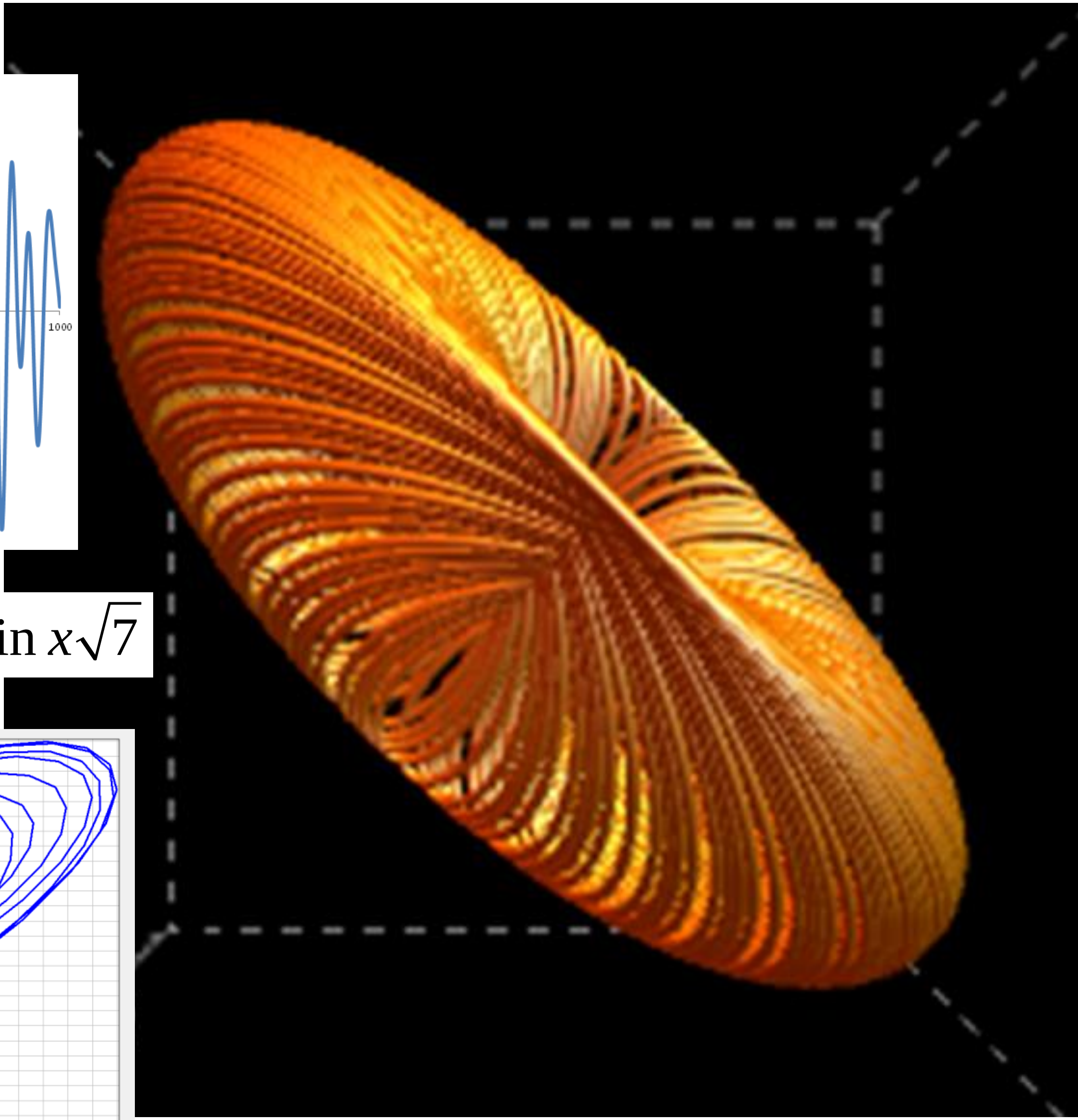
Train **Test** **Predict**

Quasiperiodic signal



$$y = \sin x + \sin x\sqrt{2} + \sin x\sqrt{5} + \sin x\sqrt{7}$$

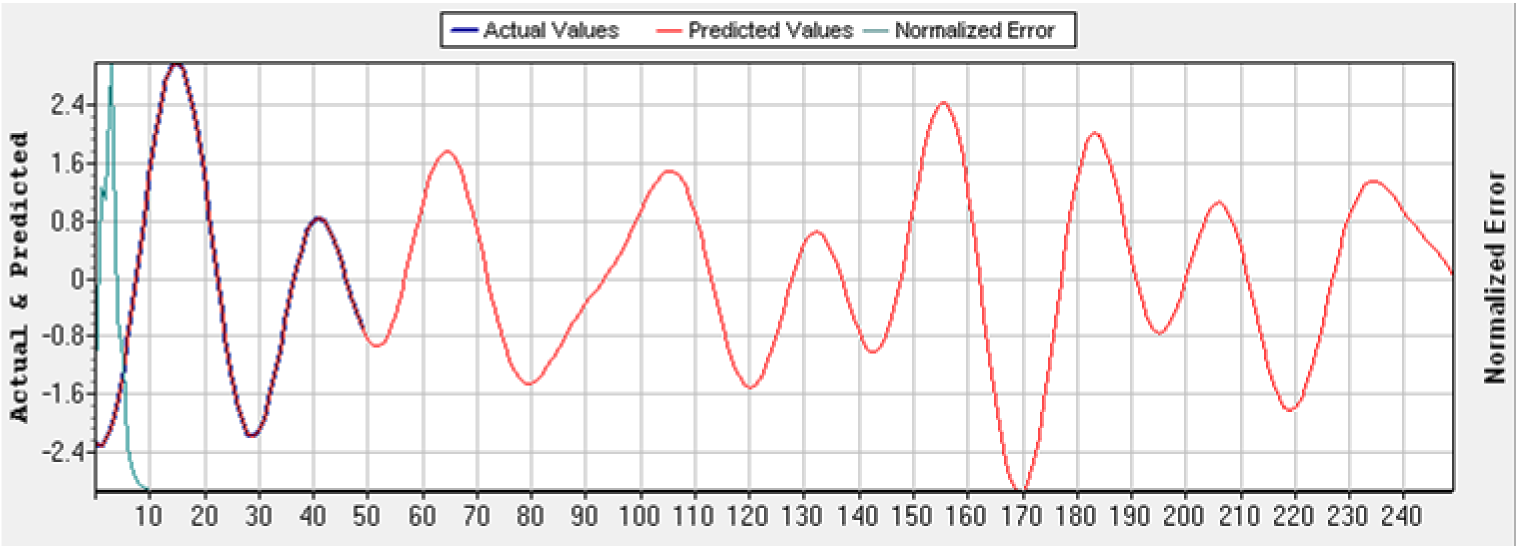


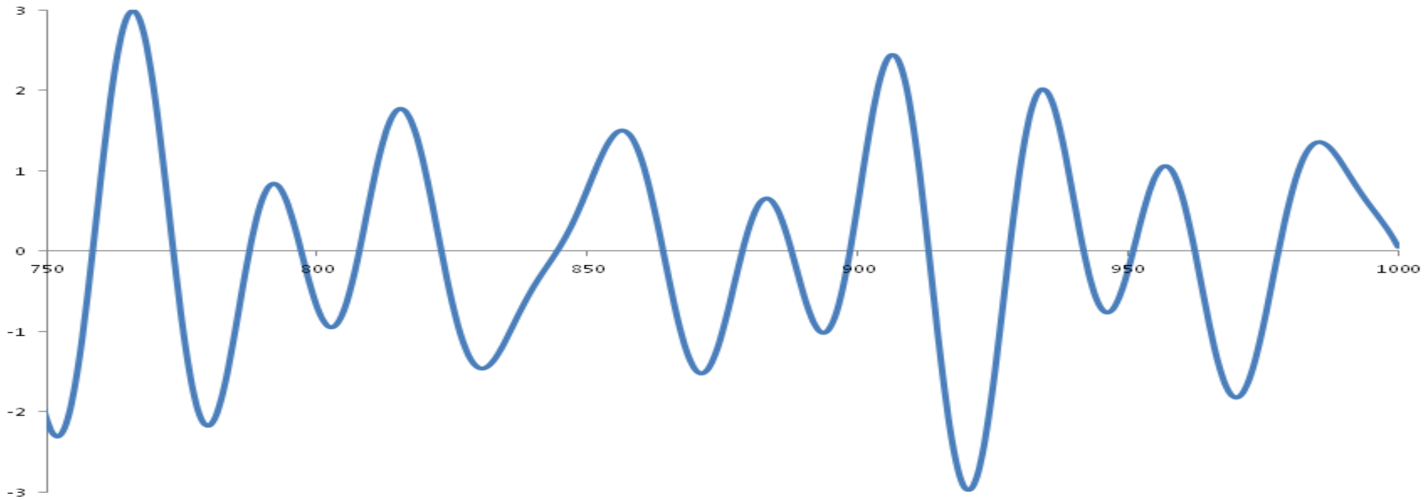


3D attractor

Quasiperiodic signal

Prediction

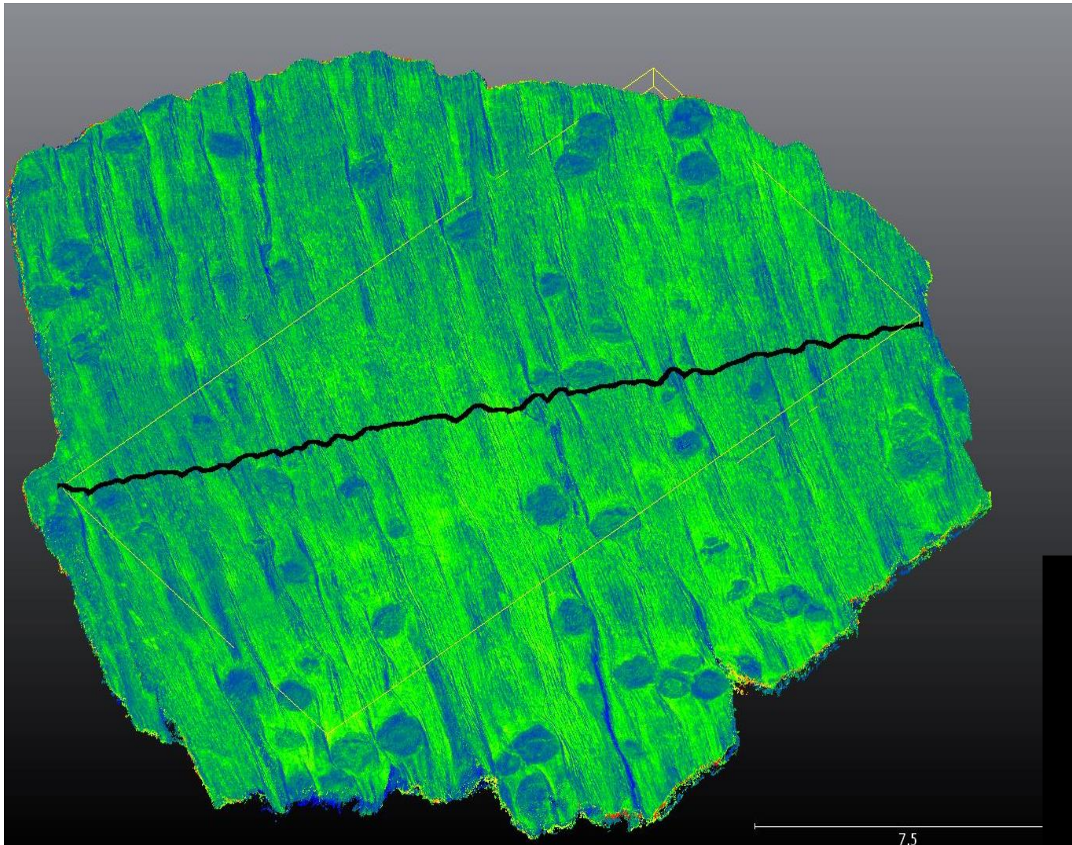




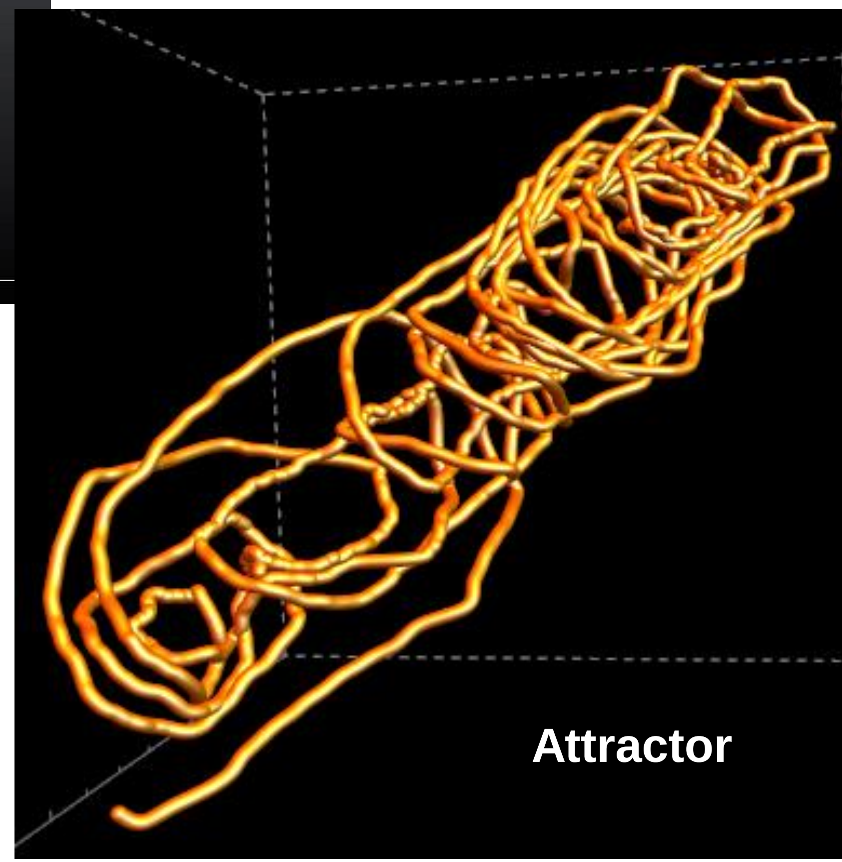
Calculated signal

$$y = \sin x + \sin x\sqrt{2} + \sin x\sqrt{5} + \sin x\sqrt{7}$$

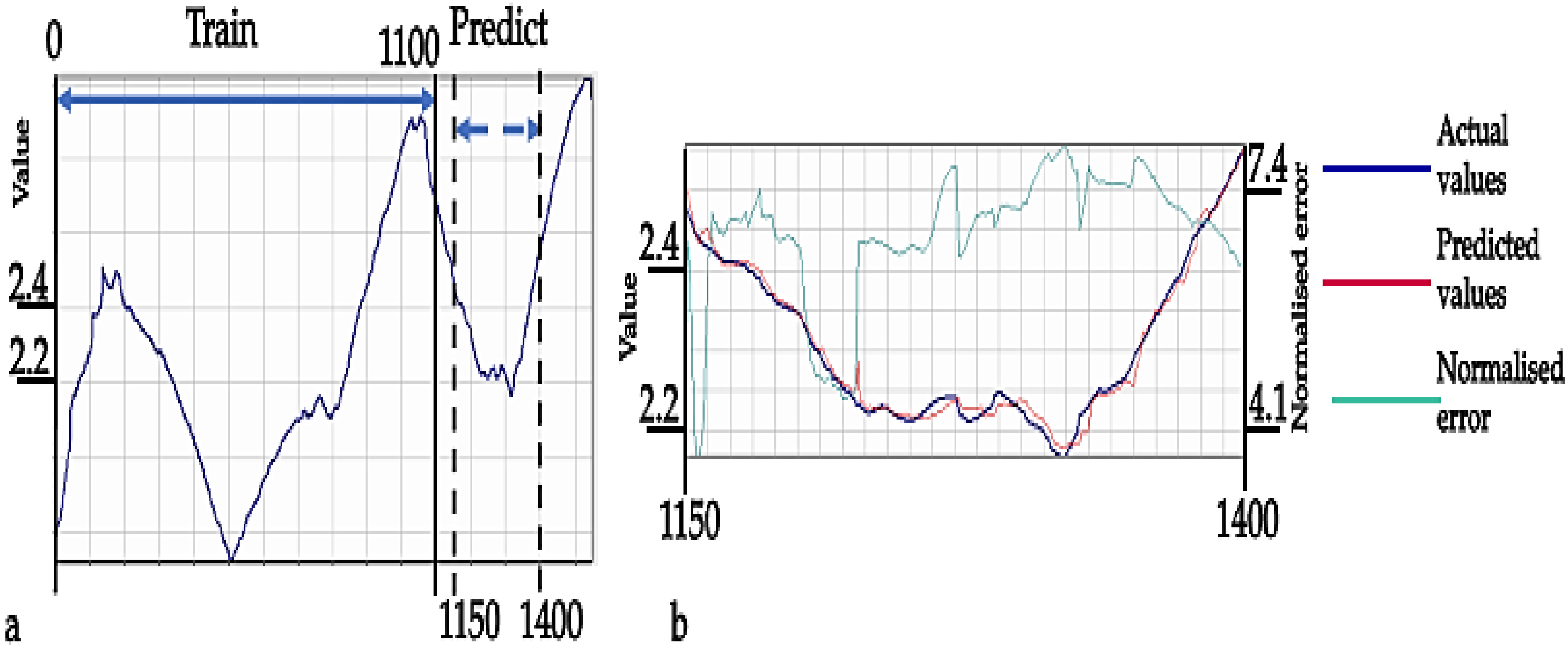
Structural Geology Prediction.

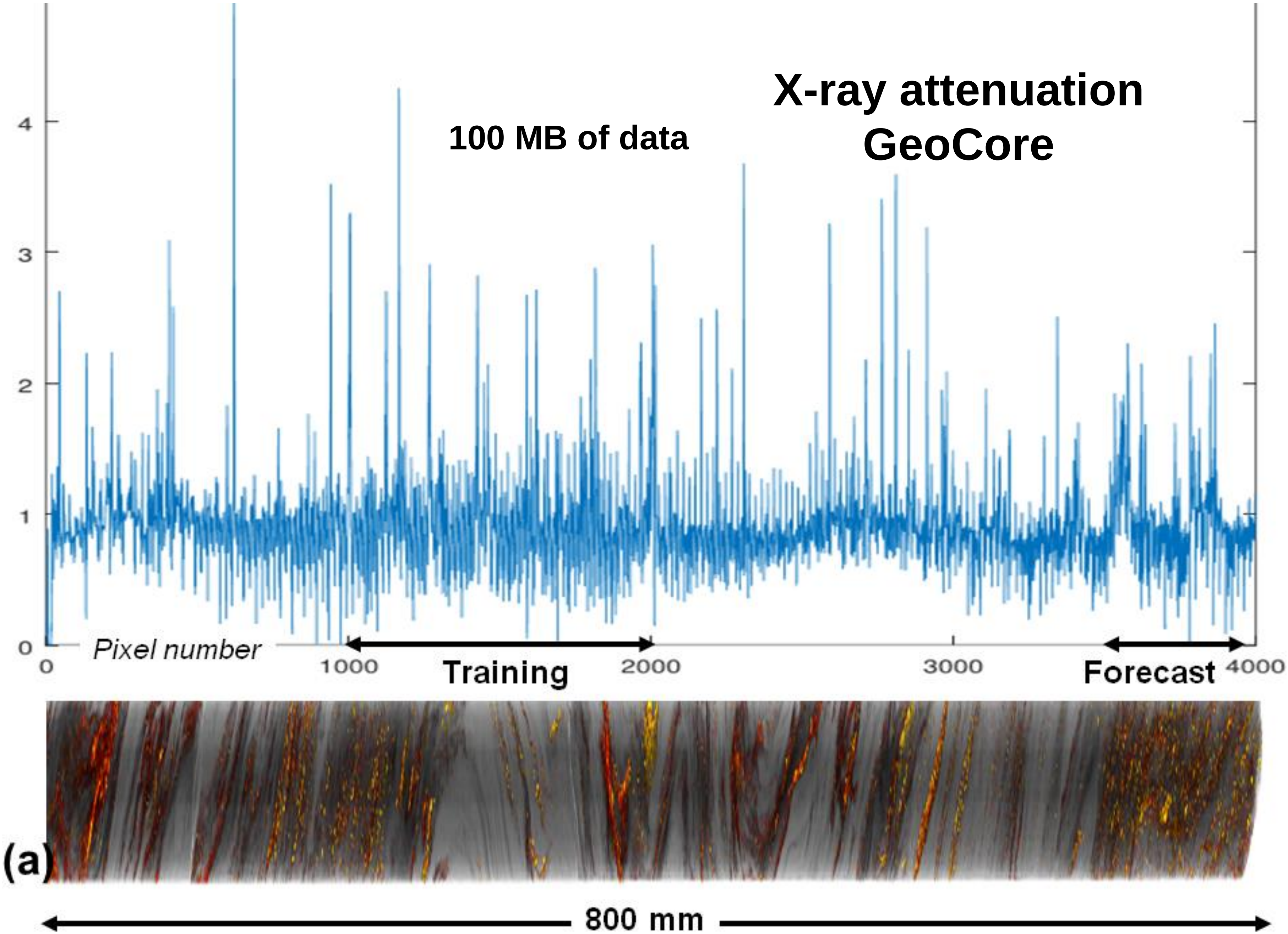


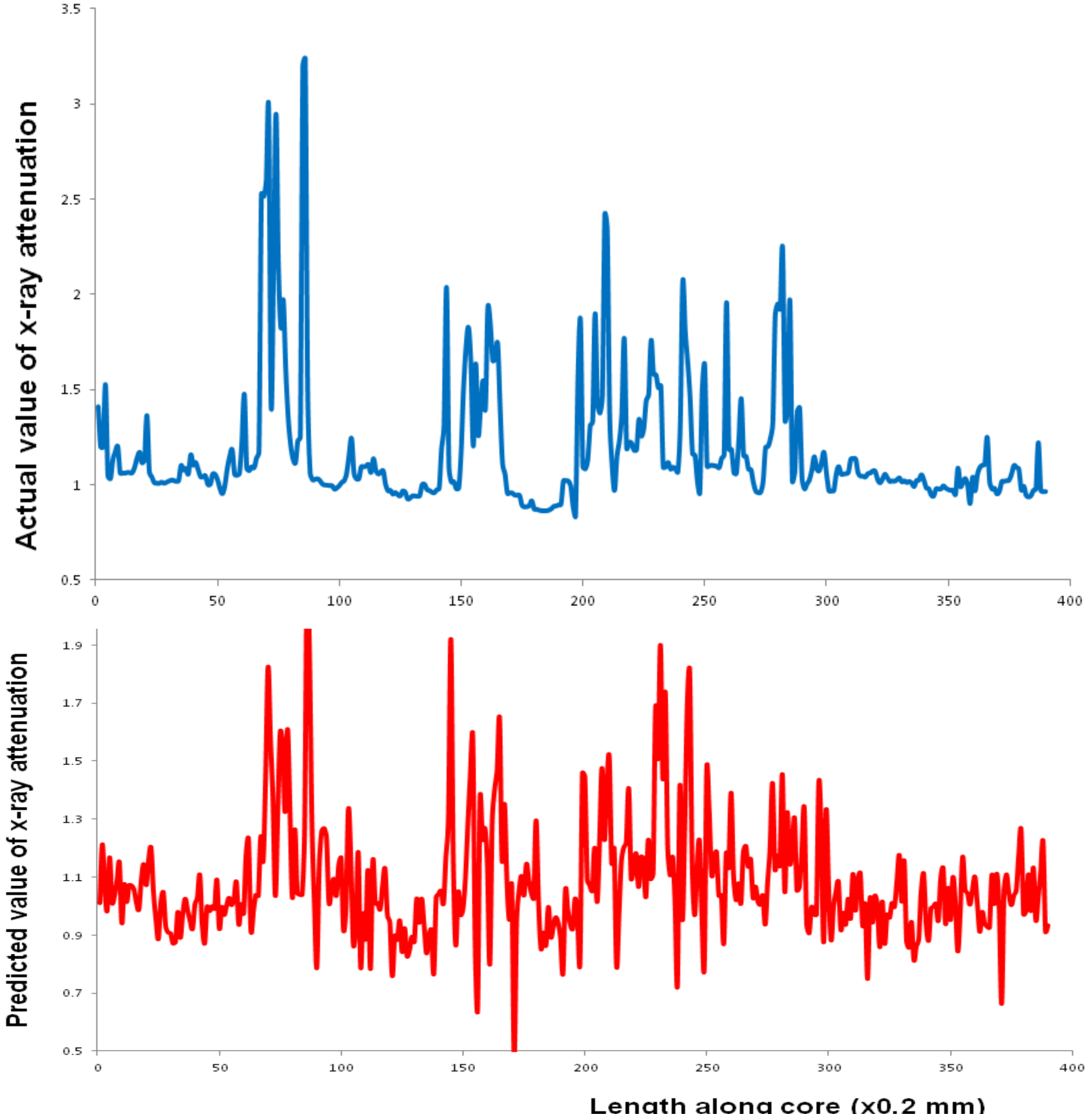
Point cloud scanned image, crenulated schist (Rum Jungle) with porphyroblasts. Profile marked in black.



PREDICTION OF FOLD PROFILE FROM BROKEN HILL




X-ray attenuation
GeoCore
100 MB of data
0
1
2
3
4
Pixel number
0
1000
2000
3000
4000
Training
Forecast
(a)
800 mm

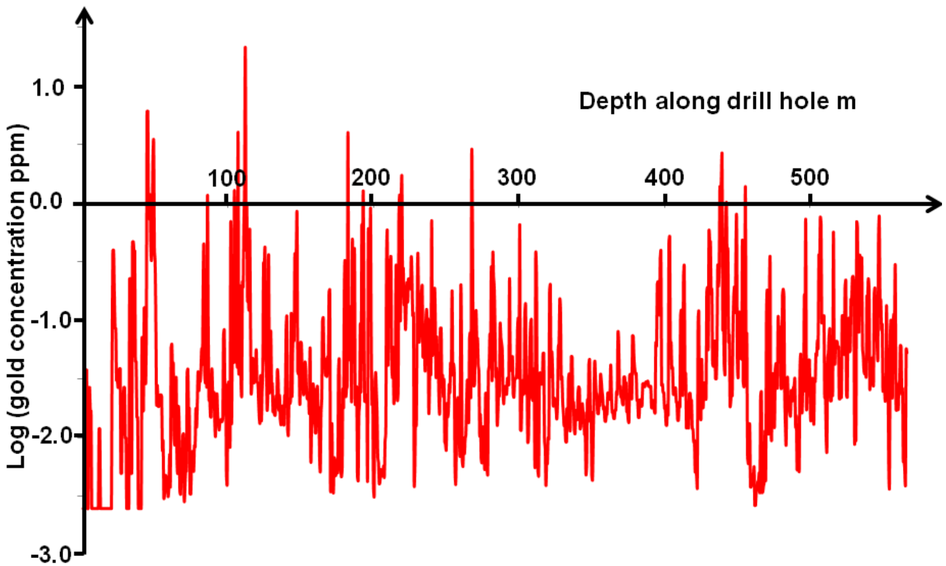


(b)

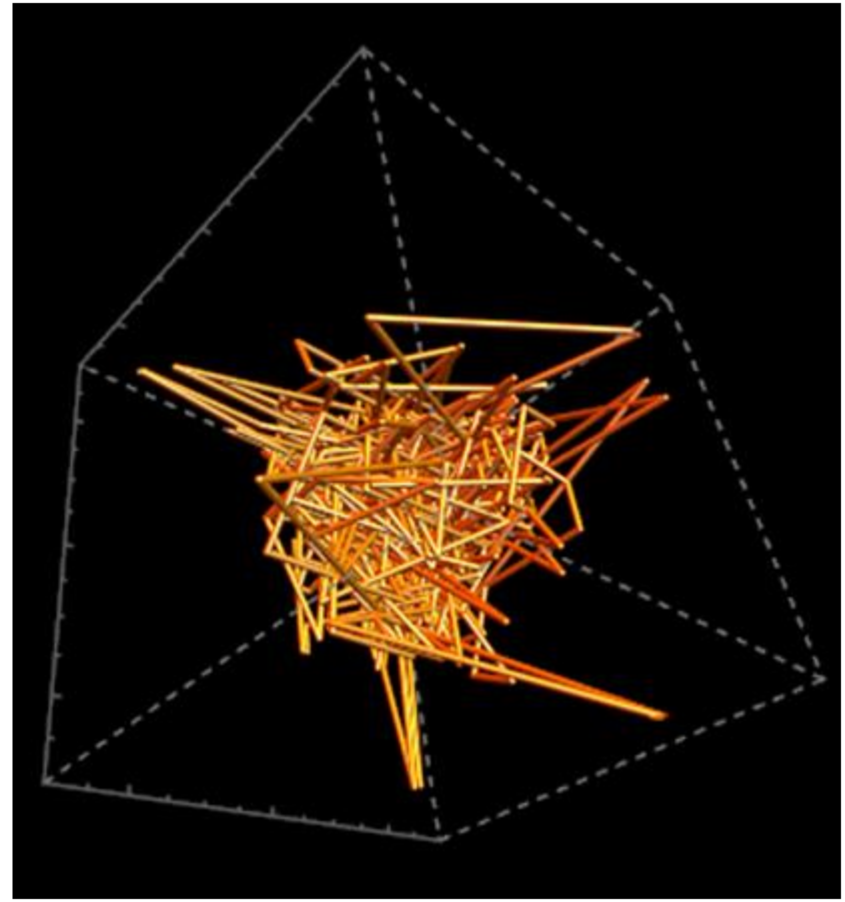
Actual vs prediction X-ray attenuation

GeoCore

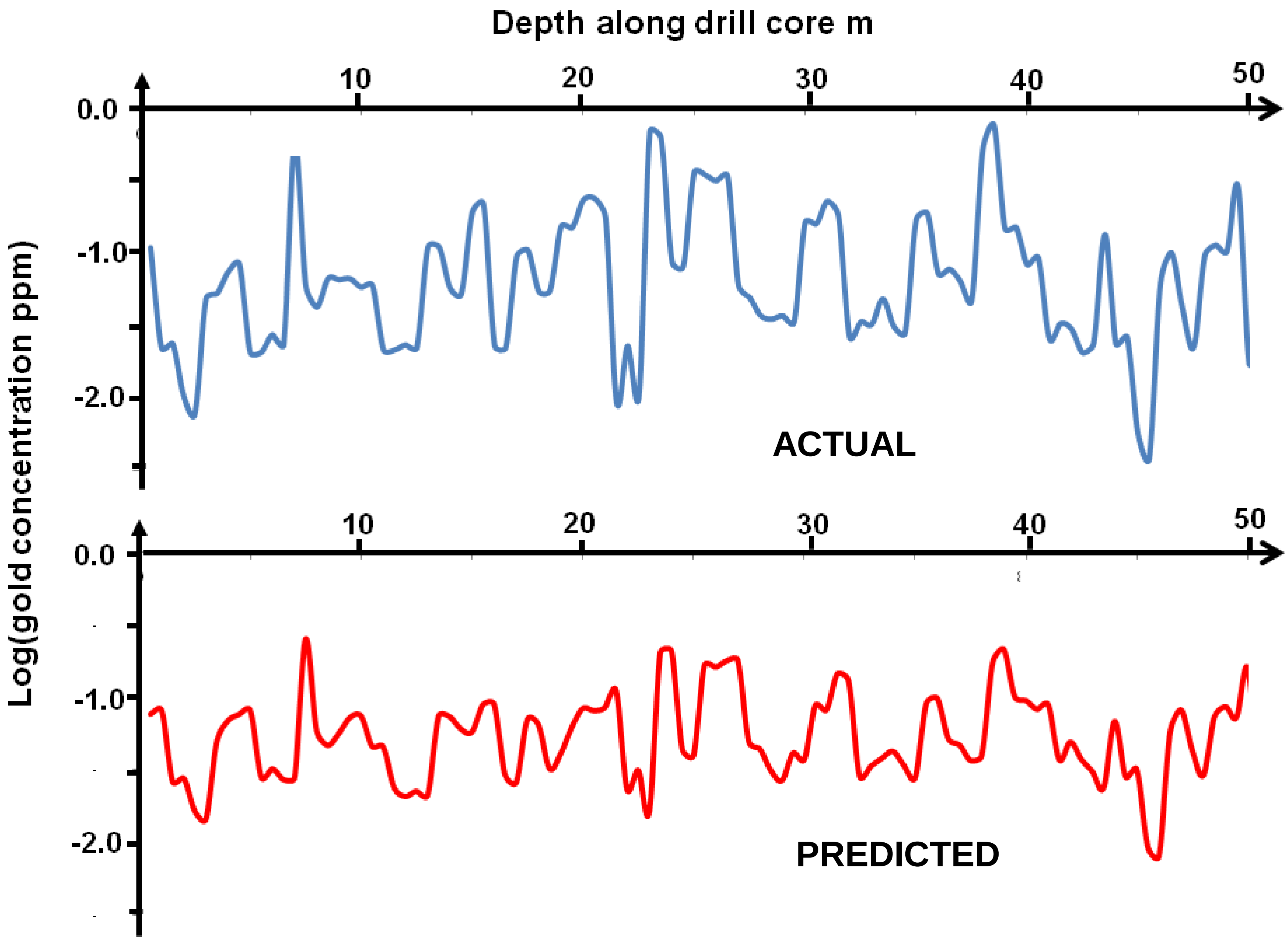
PREDICTION OF NATURAL GOLD CONCENTRATION

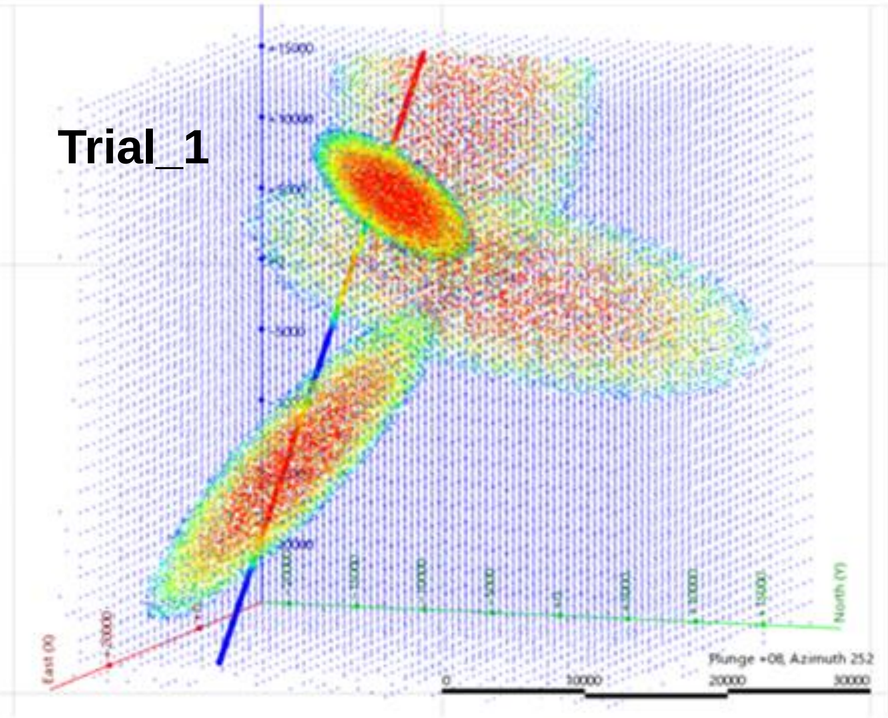


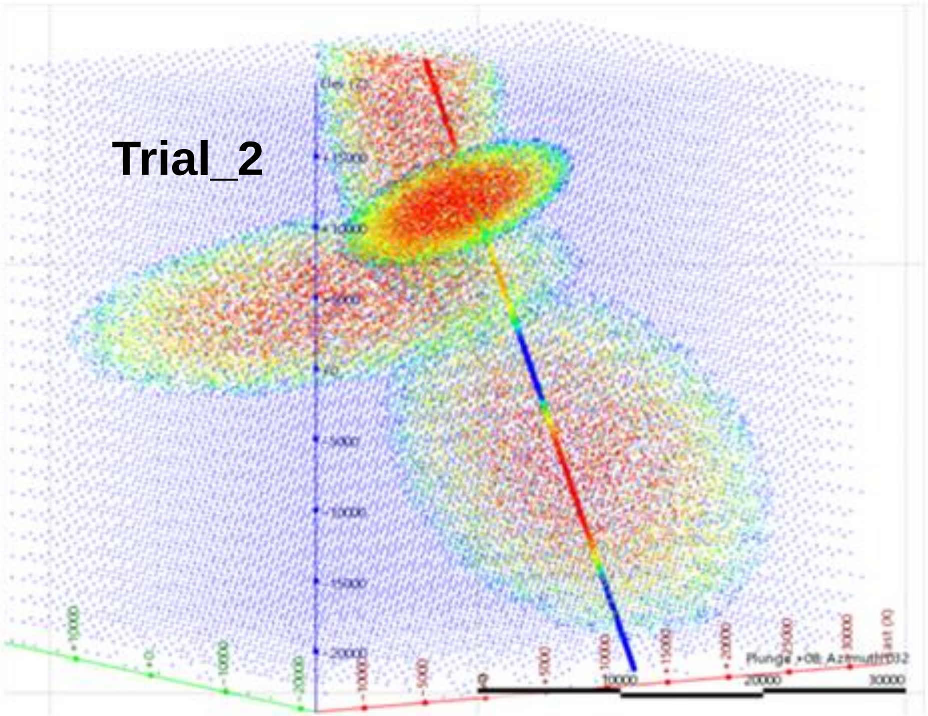
Log (gold concentration ppm) from Yilgarn gold deposit

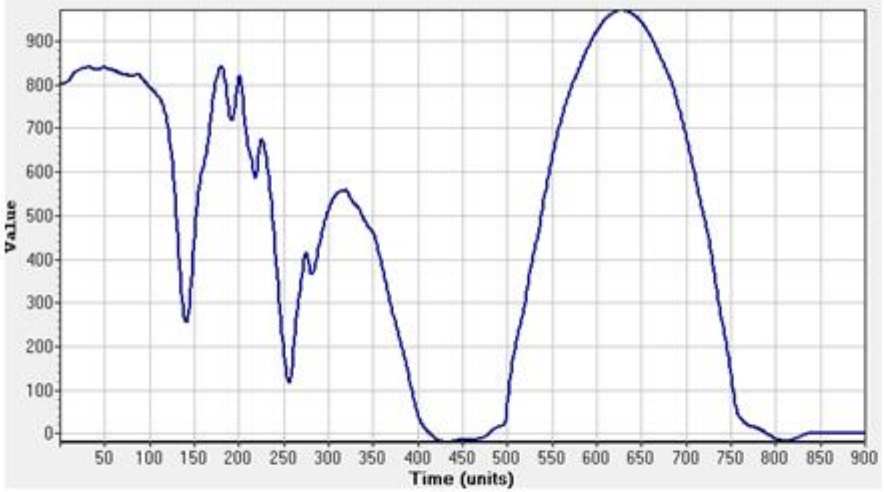


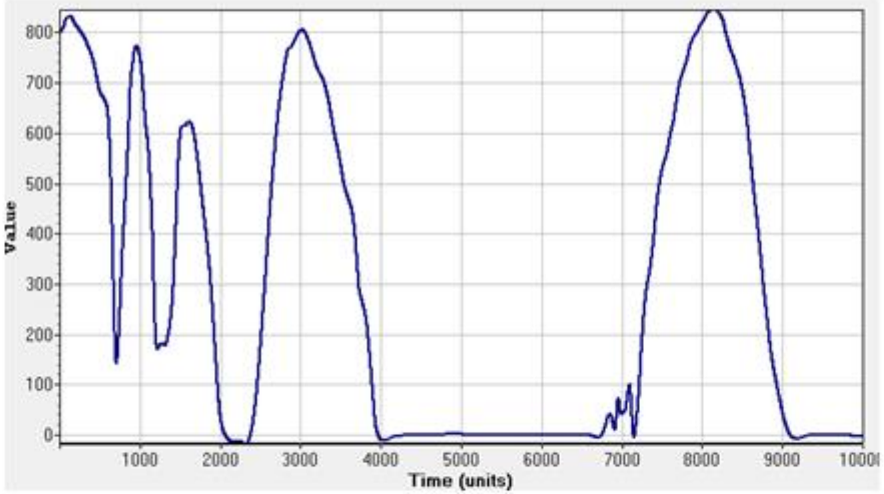
3D attractor projected from ≈10 dimensions


Depth along drill core m
10
20
30
40
50
0.0
-1.0
-2.0
ACTUAL
10
20
30
40
50
0.0
-1.0
-2.0
PREDICTED
Log(gold concentration ppm)

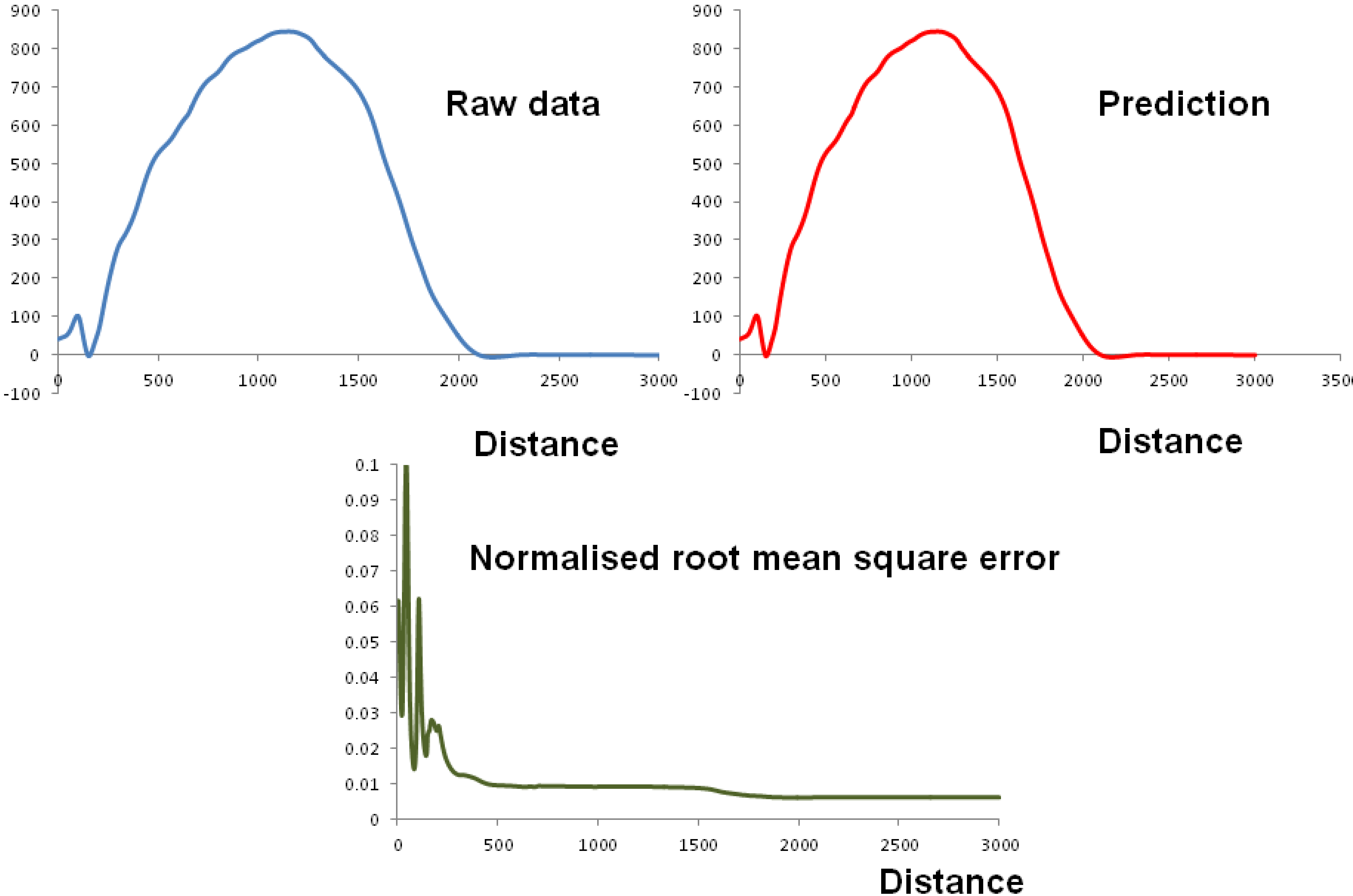


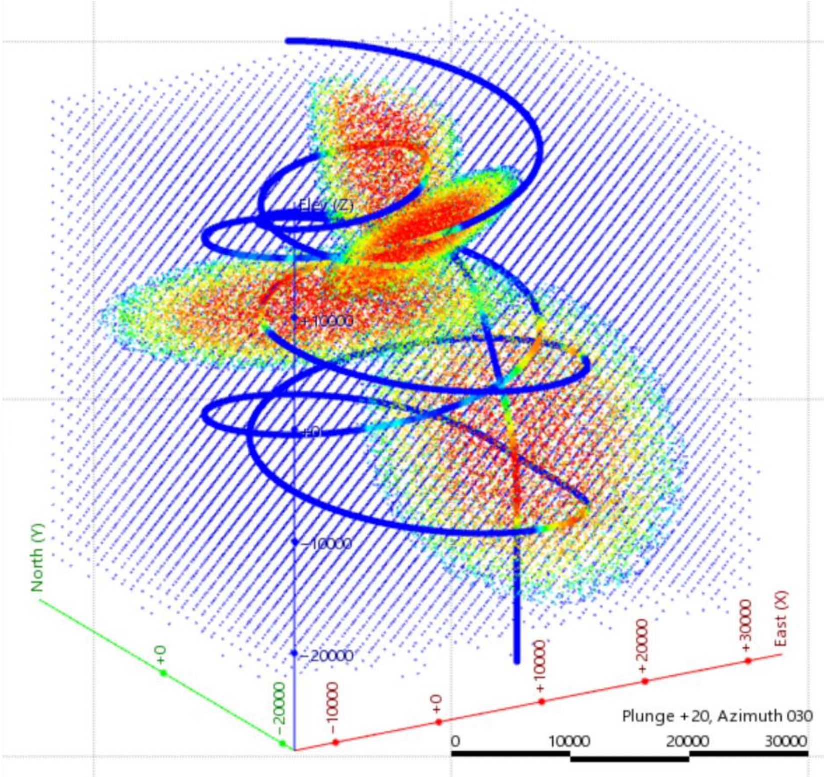


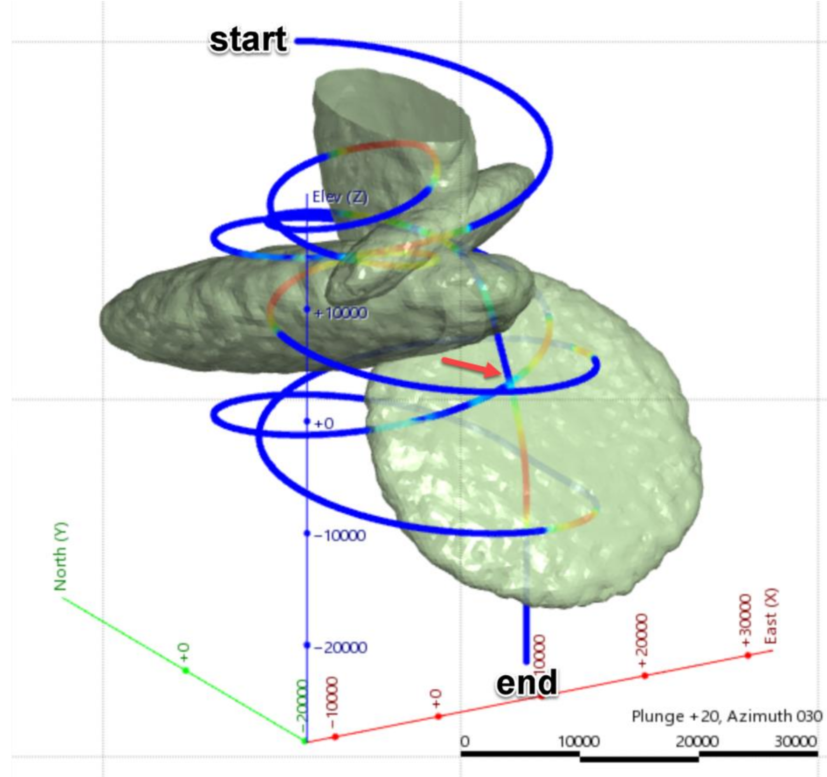




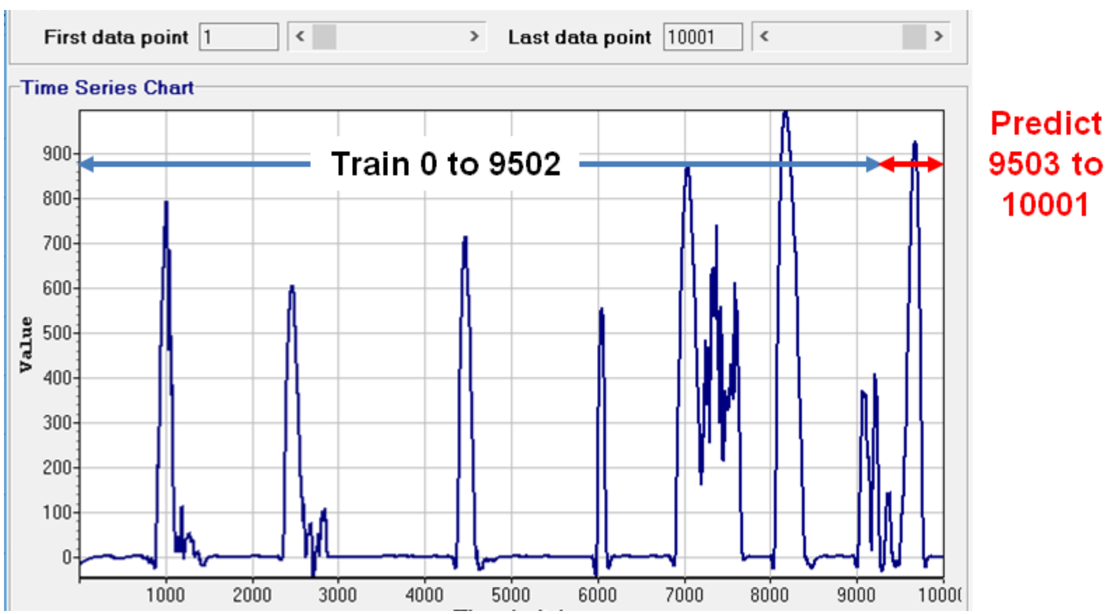
Models constructed by Jun Cowan


Raw data
Distance
Prediction
Distance
Normalised root mean square error
Distance
900
800
700
600
500
400
300
200
100
0
-100
0
500
1000
1500
2000
2500
3000
3500
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01

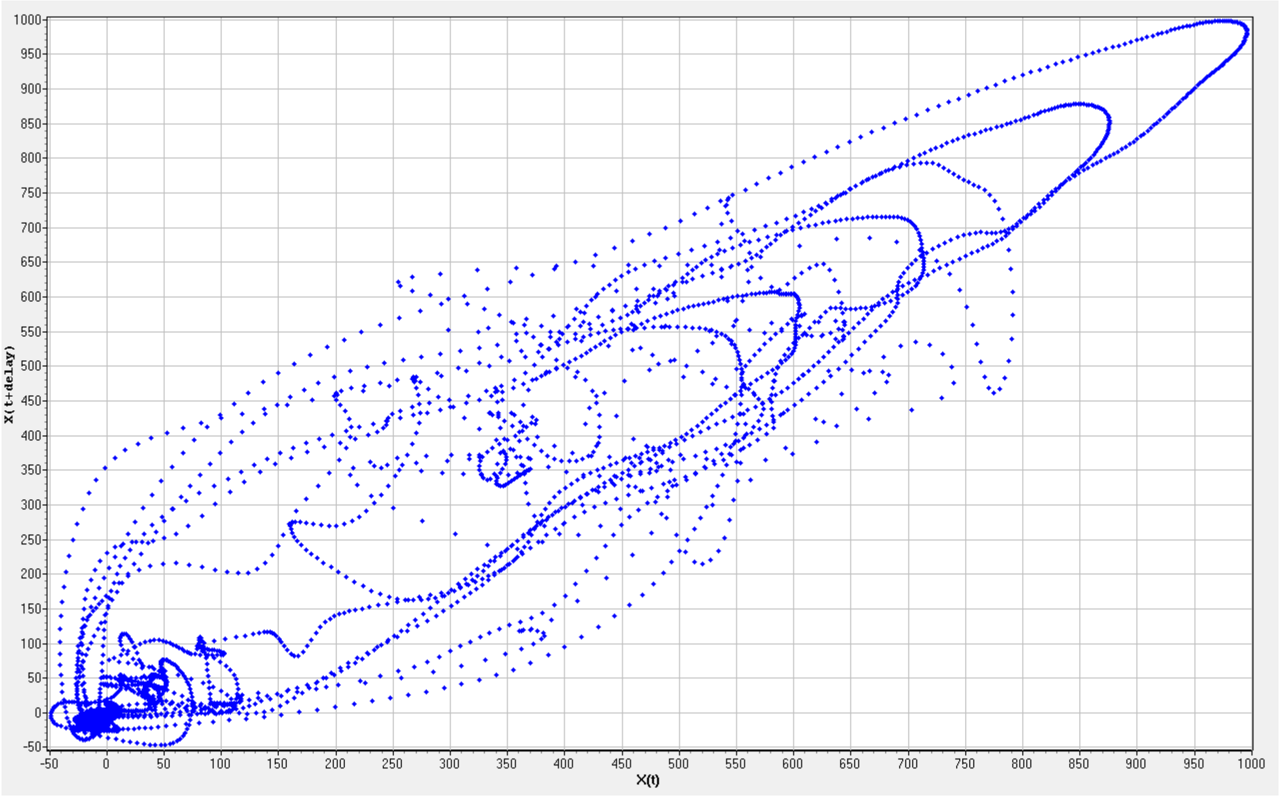




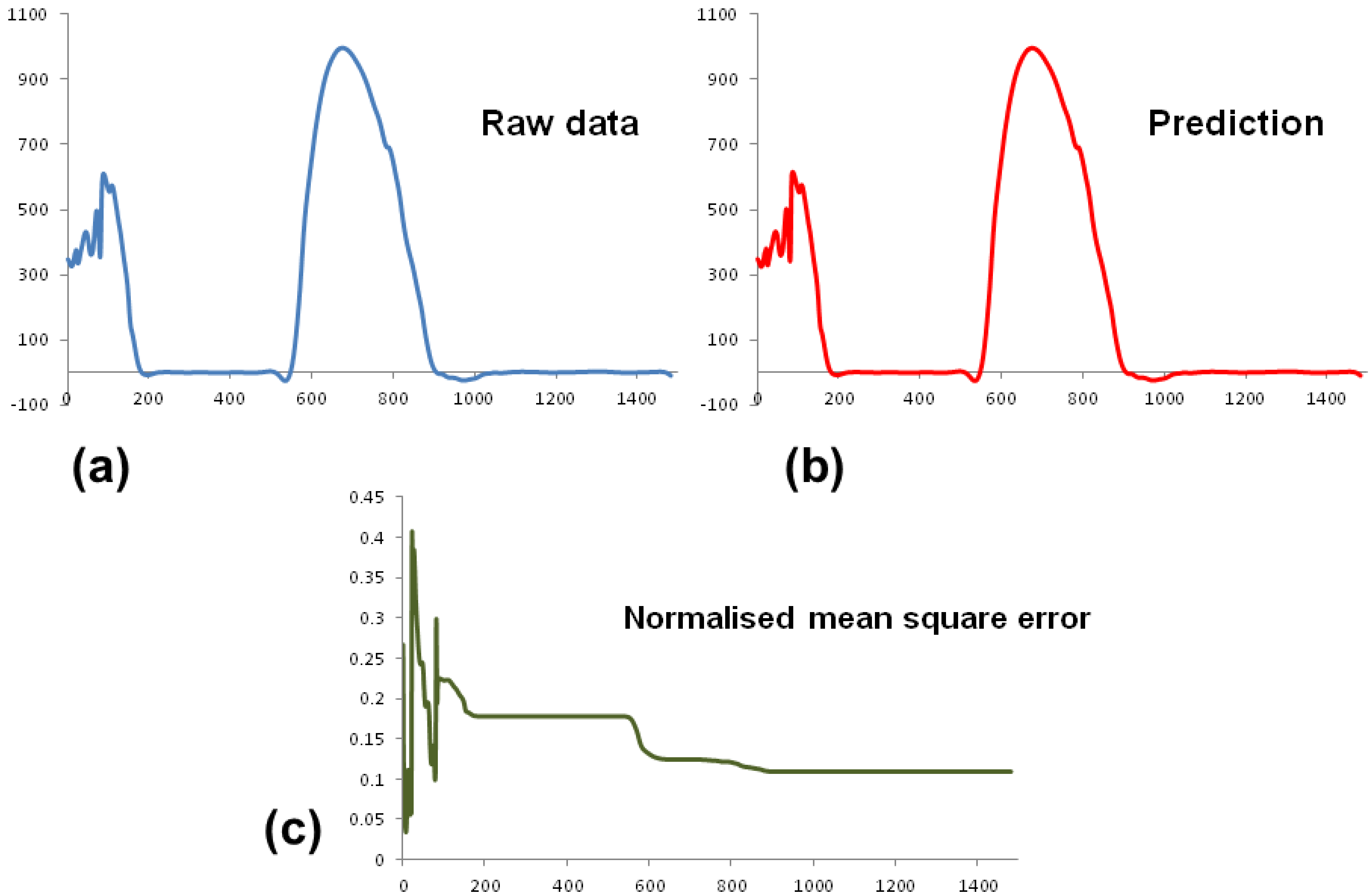
Trial_3



Models constructed by Jun Cowan



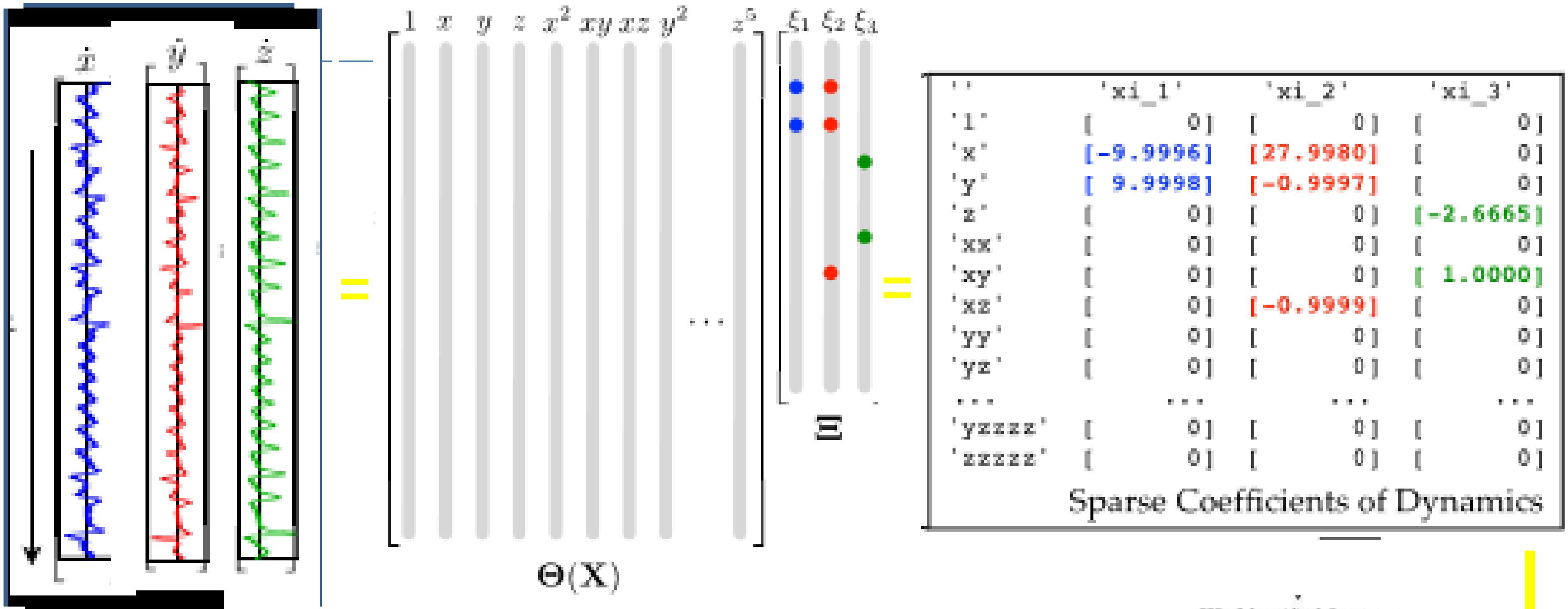
2D attractor for Trial_3


Raw data
1100
900
700
500
300
100
-100
0
200
400
600
800
1000
1200
1400
(a)
Prediction
(b)
Normalised mean square error
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
(c)

Trial_3

An exciting new development that surpasses all attractor based prediction methods.

DATA DRIVEN SPARSE IDENTIFICATION
OF
NONLINEAR DYNAMICAL SYSTEMS

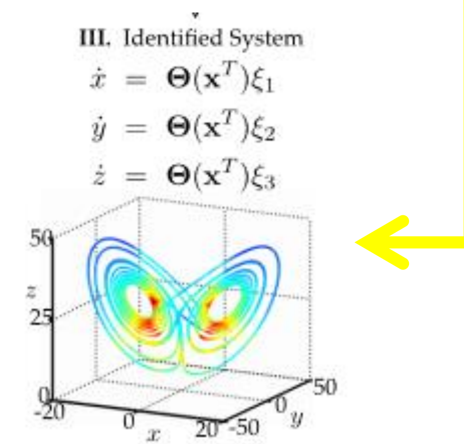


''	'xi_1'	'xi_2'	'xi_3'
'1'	[0]	[0]	[0]
'x'	[-9.9996]	[27.9980]	[0]
'y'	[9.9998]	[-0.9997]	[0]
'z'	[0]	[0]	[-2.6665]
'xx'	[0]	[0]	[0]
'xy'	[0]	[0]	[1.0000]
'xz'	[0]	[-0.9999]	[0]
'yy'	[0]	[0]	[0]
'yz'	[0]	[0]	[0]
...	...	...	...
'yzzzz'	[0]	[0]	[0]
'zzzzz'	[0]	[0]	[0]

Data

Library of functions

Coefficients for best fit assuming the simplest model



IDENTIFICATION OF SYSTEM DYNAMICS FROM DATA USING SPARSE REGRESSION

After Brunton et al., 2016

DATA DRIVEN SPARSE IDENTIFICATION OF NONLINEAR DYANMICAL SYSTEMS

In the past 2-3 years these methods have been applied in many scientific, engineering and industrial applications.

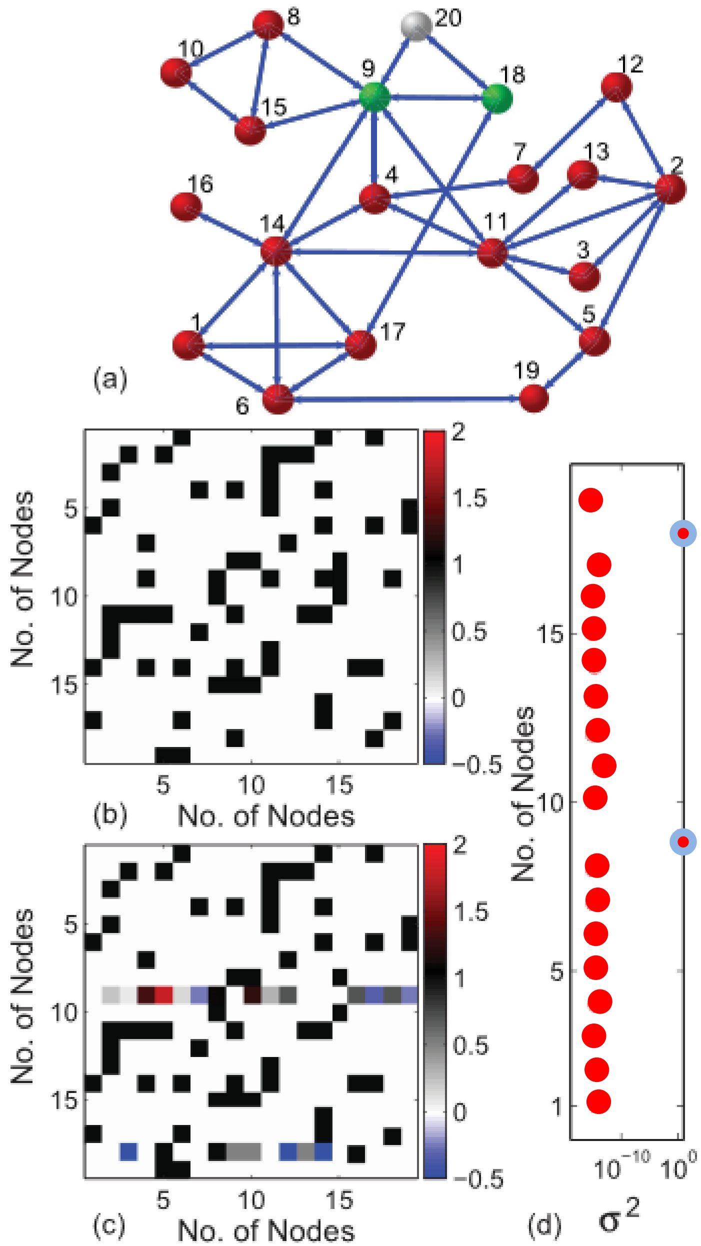
An example is in chemical engineering and drug design.

There are many chemical compounds we do not know about. If we prescribe the response we want then sparse regression is used to find the simplest assembly of atoms one needs to produce that response and hence synthesise the necessary chemical compound.

Some are claiming this is a completely new paradigm in science where data, not theory or simulation, drives new developments.

An example: Detection of hidden nodes

Network: Hidden node is #20



Actual adjacency matrix

Calculated adjacency matrix showing anomalous entries for nodes 9 and 18

Variance for calculated coefficients showing anomalous values for nodes 9 and 18

***From Su et al., Phys. Rev., E85, 2012*.**

These nonparametric approaches, driven solely by the data, represent entirely new ways of prediction, analysing, interpreting and thinking about geoscientific data.

They offer ways of predicting and inverting structural, geophysical and geochemical data sets with no ambiguity.

They even offer ways of detecting unknown components of a system given only the data with no underlying statistical or model assumptions.

The message is:

Understanding is in the data

NOT in the statistics (kriging, potential field inversion) which are always ambiguous

Knowledge is in the dynamics

Which are derived from nonlinear analysis

The results of which are sparse, parsimonious and unique.

How much data do we need?

We can collect N data sets (Terabytes) and it is expensive.

But only K data sets are needed to define a solution with K very much smaller than N.

If we collect 2K data sets at random the problem is solved.

K is revealed by the dynamics of the processes involved.

These non-parametric, nonlinear approaches and their rapidly developing generalisations represent the future of data analytics in the geosciences.

Thank you and have fun.