Complexity Methods for Behavioural Science

Basic Timeseries Analysis
Basic Nonlinear Timeseries Analysis

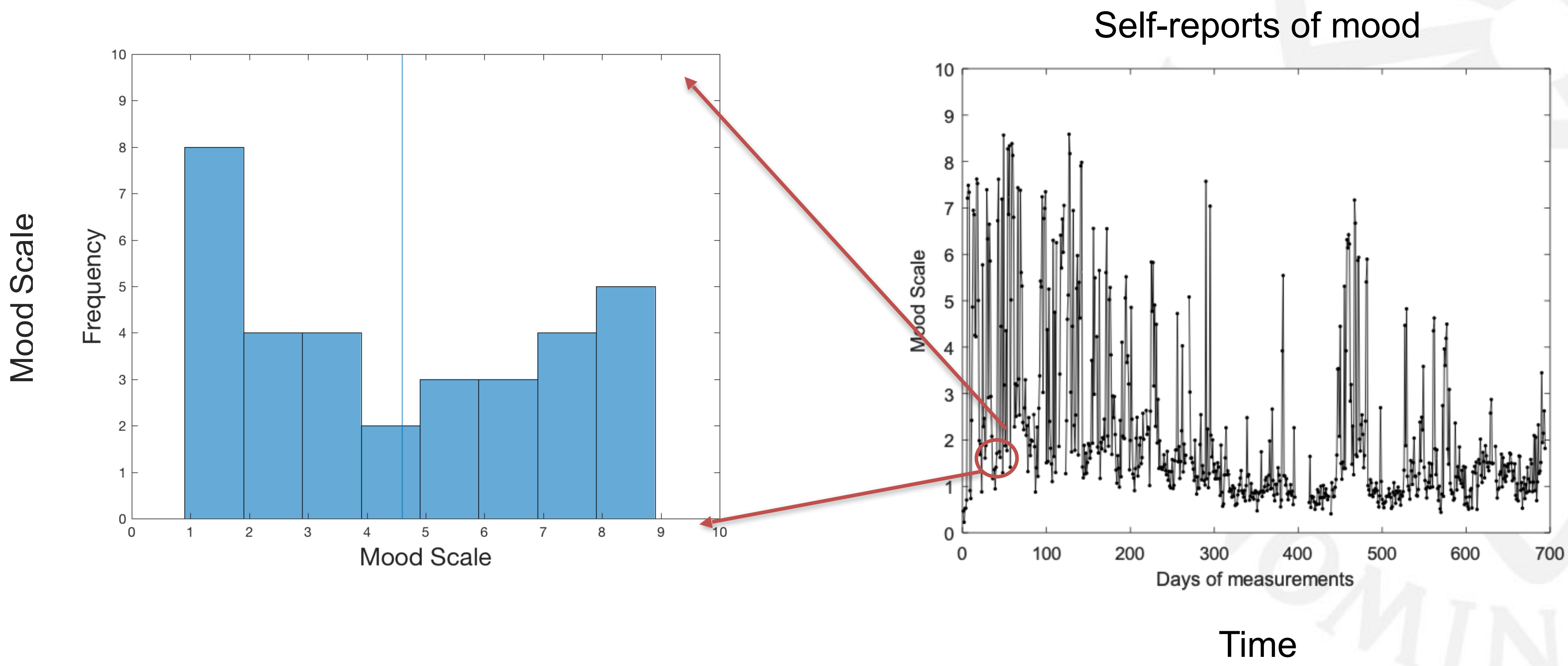
Scaling

f.hasselman@bsi.ru.nl

Behavioural Science Institute
Radboud University Nijmegen

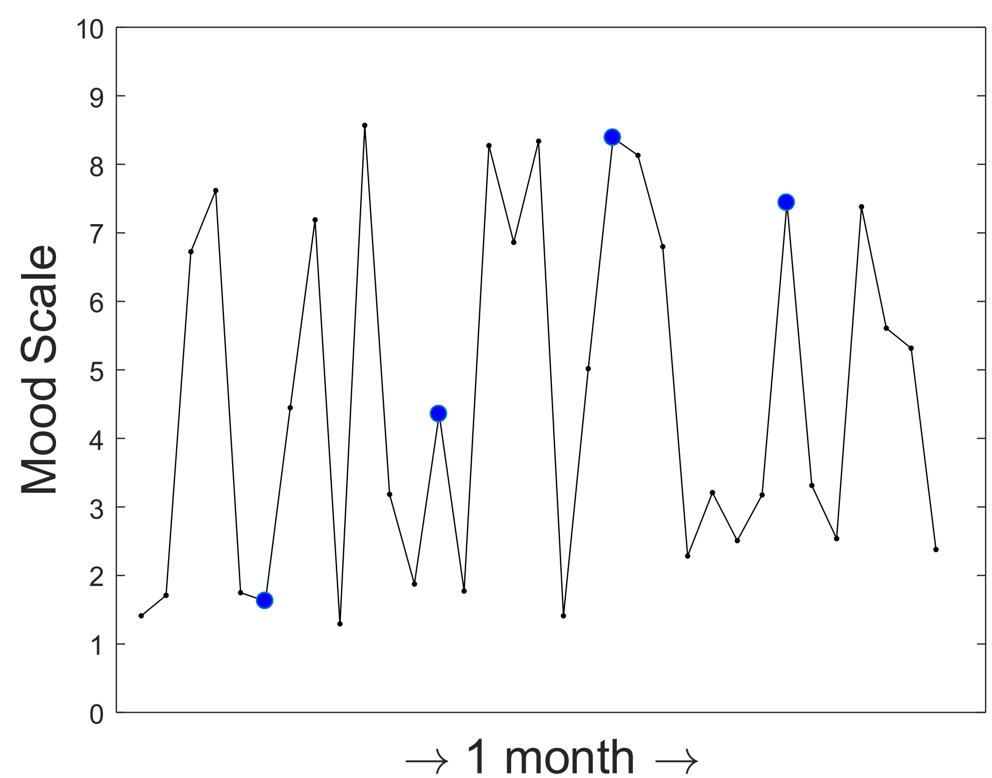
Depression as a dynamic process

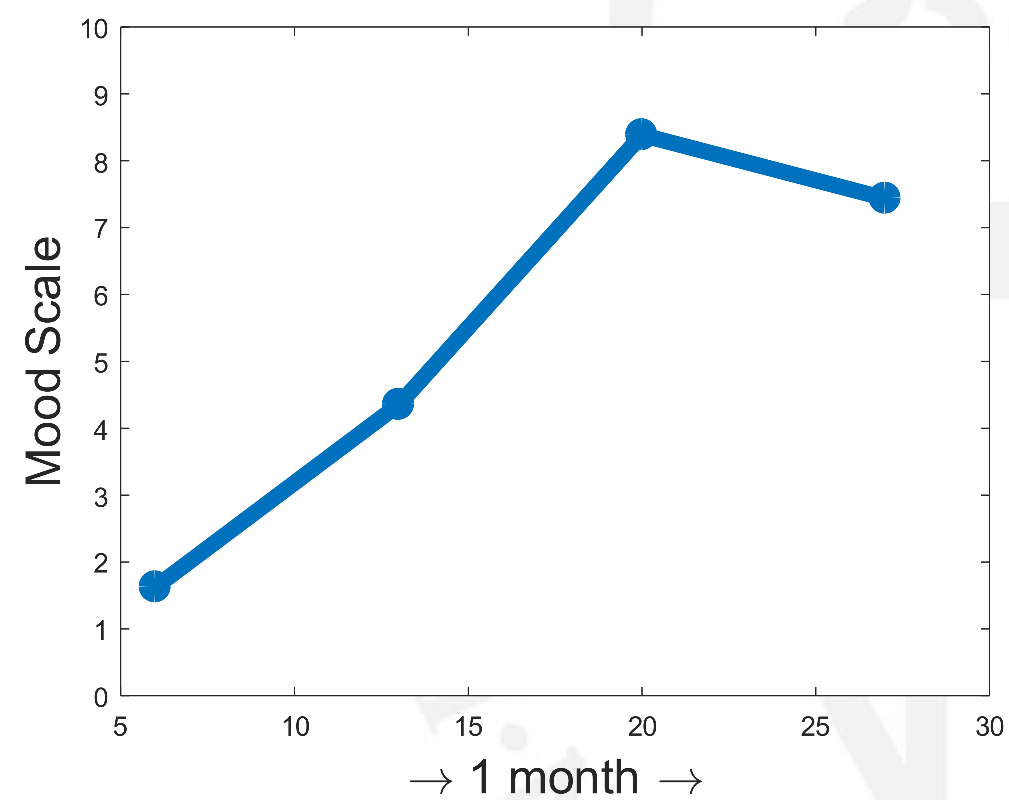
Depression is :
- By definition a dynamic change process
- Dynamics are lost in standard analyses



Depression as a dynamic process: #SocialSaturday

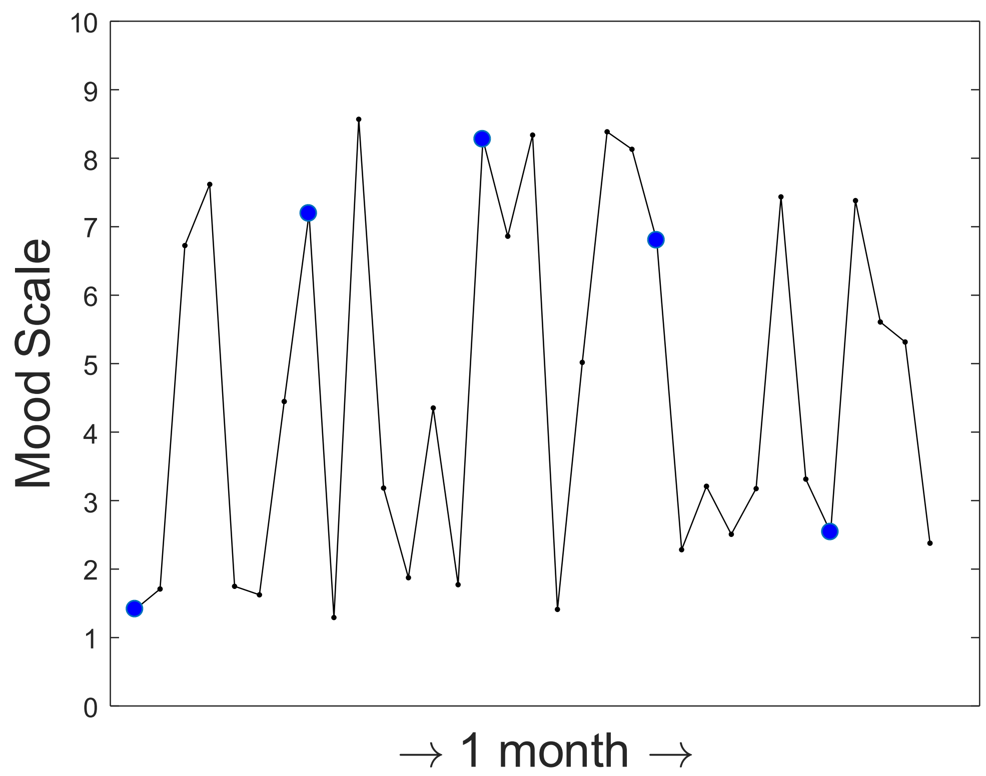
Why collecting time series is useful

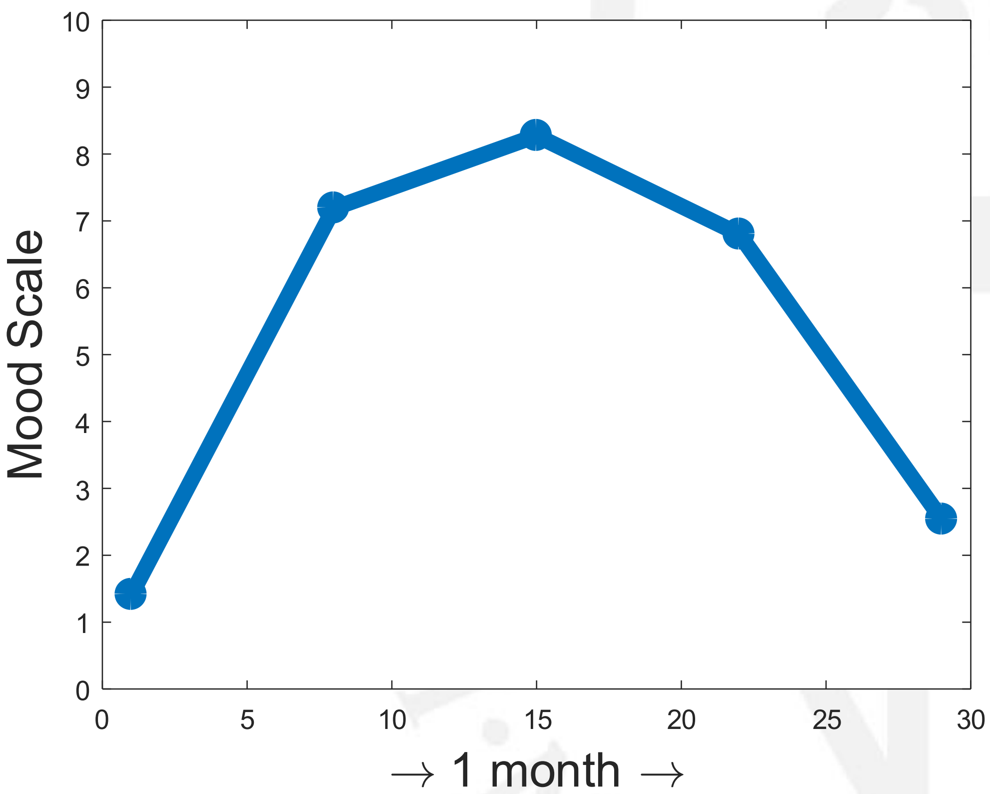




Depression as a dynamic process: #MaxoutMonday

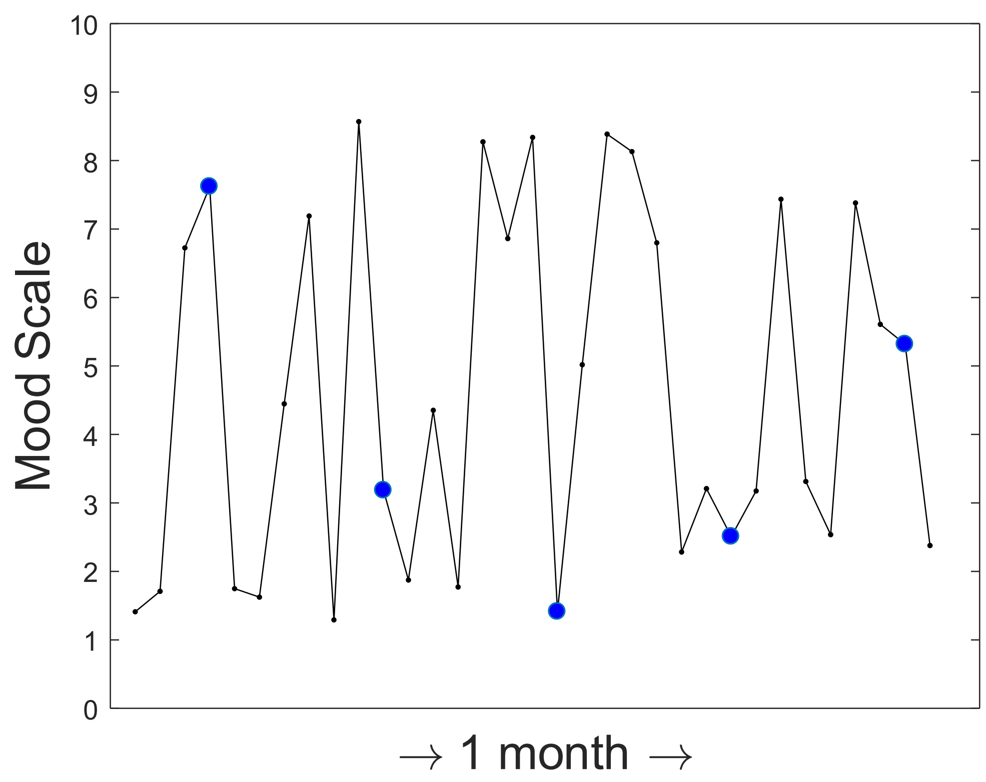
Why collecting time series is useful

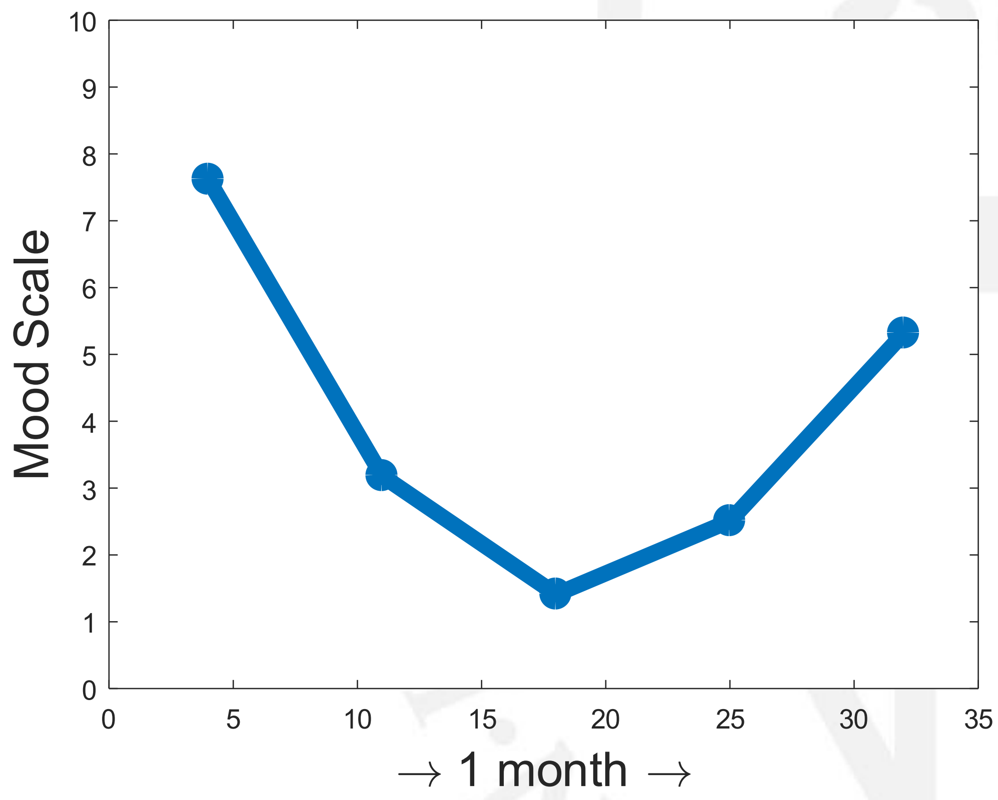




Depression as a dynamic process: #ThursdayThoughts

Why collecting time series is useful

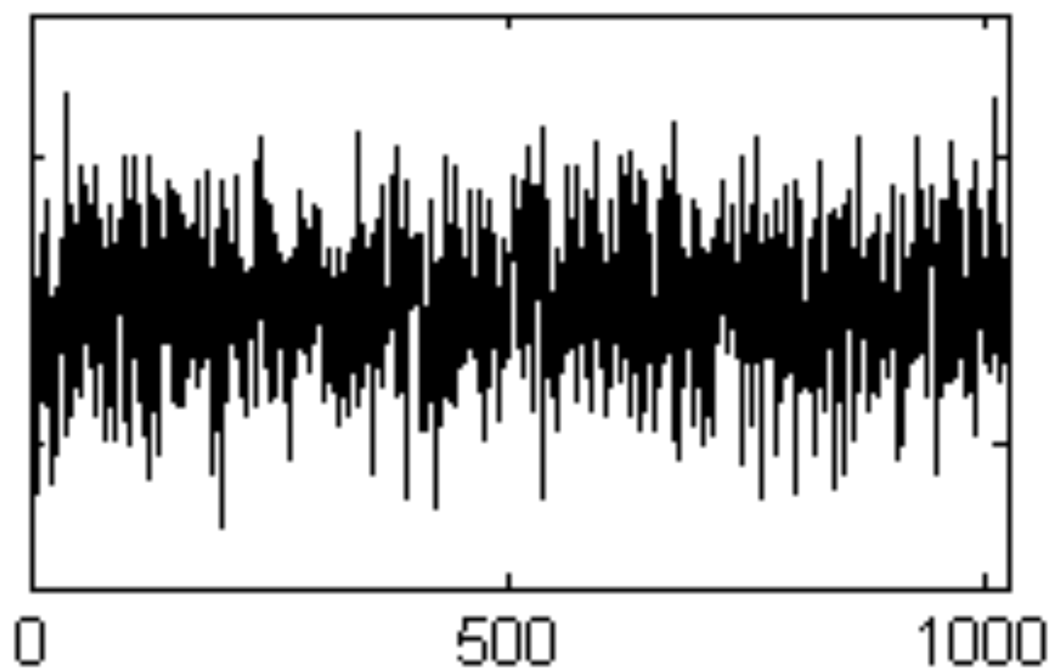




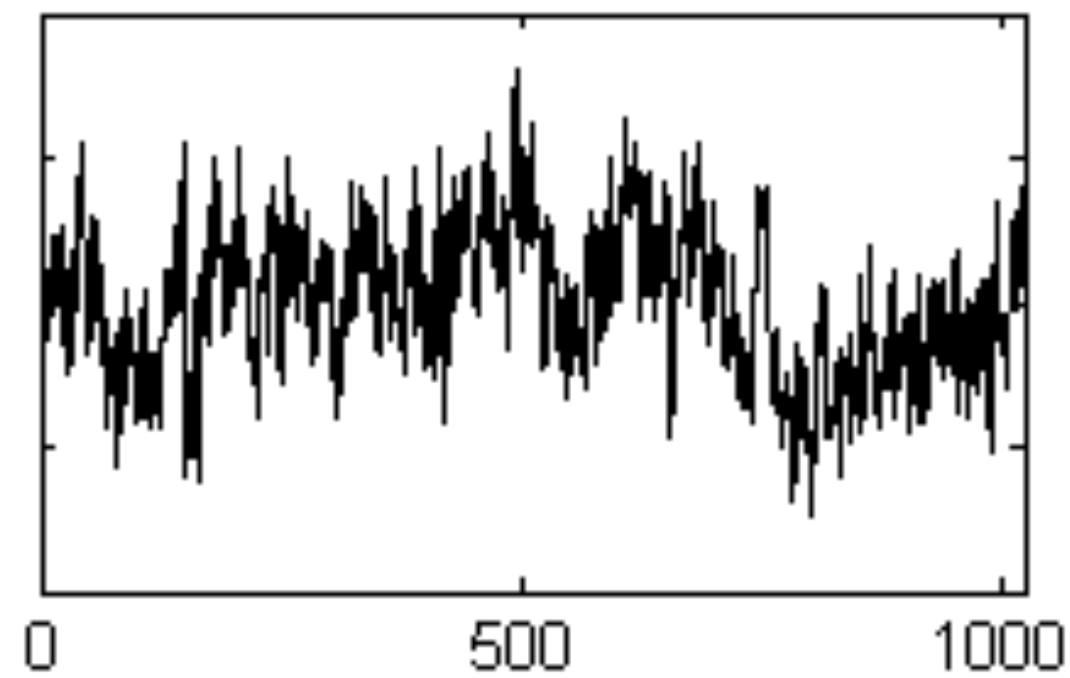
Alternative: Analyse then (perhaps) aggregate

Collect many responses over time

- Keep the order of the data points intact
- Observe temporal structure of variability
- → How does the process change over time?



Random variability



Structured variability

➔ How to distinguish random and non-random behavior?

Story so far …

Deterministic mathematical models of change processes:

- Simulate the temporal-evolution of one or more system observables under different values of a control parameter and/or resource levels and/or coupling
- Can reveal temporal patterns that resemble a random process
- Show interaction dynamics in an N-dimensional state space when N processes are dynamically coupled or mutually dependent

- *Background:* Complexity Science / Science of the Individual / Systems Biology

Today:

- Random events can reveal temporal patterns that resemble a deterministic process
- Analysing temporal correlations (correlation functions ~~and AR-MA models~~)
- Basic quantification of temporal patterns (ACF/PACF/CCF, Relative Roughness,
- Entropy)

Correlation Functions

State space

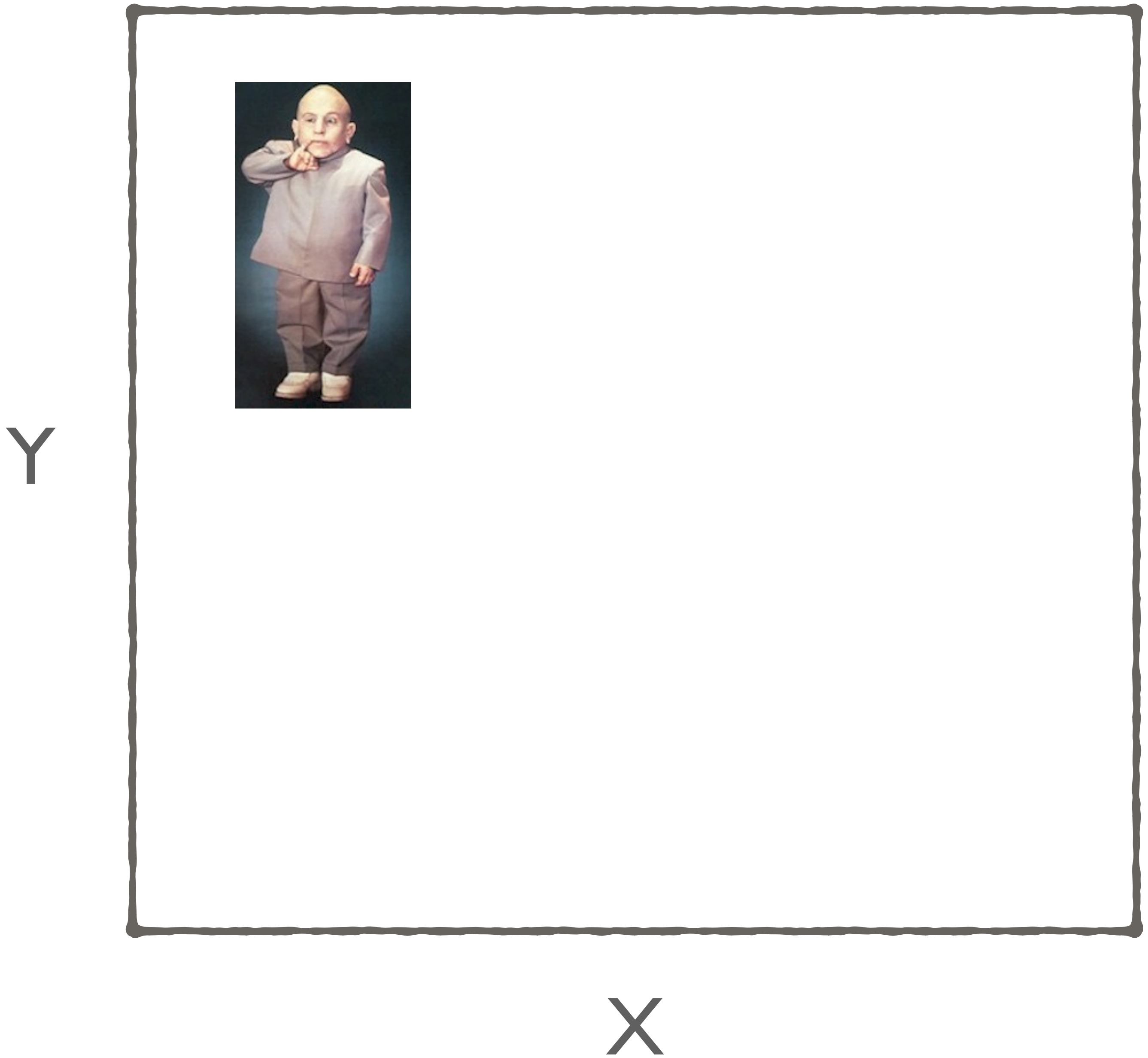
Time scales

Linear v. nonlinear

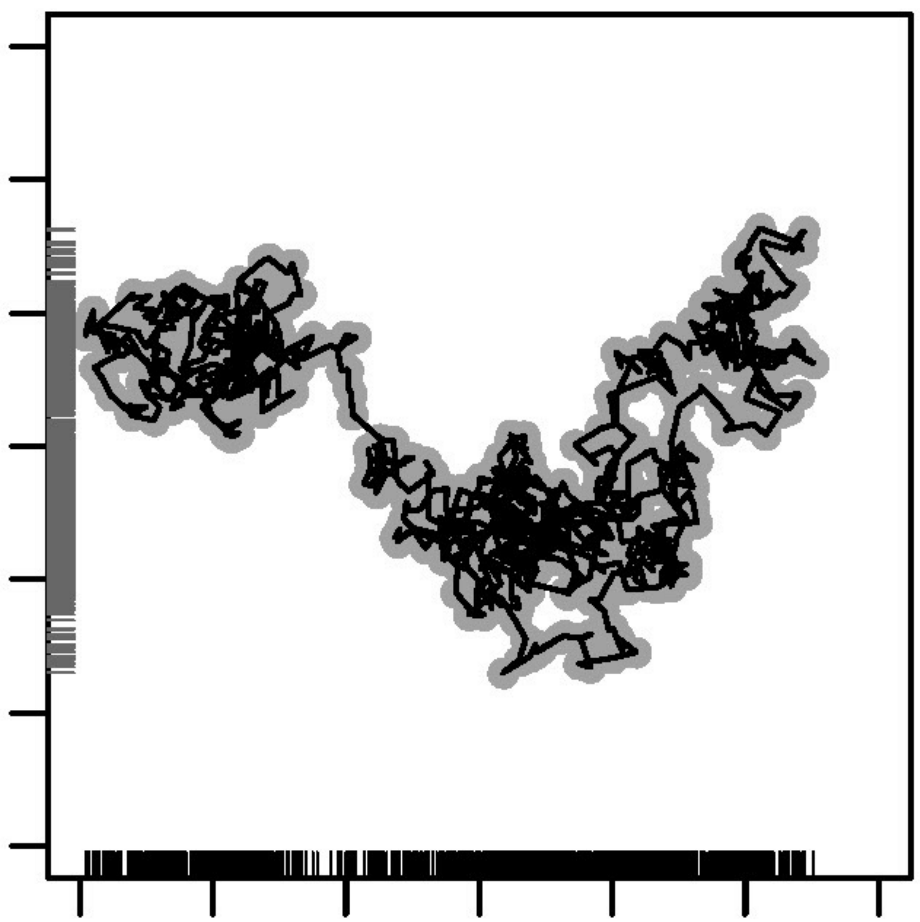
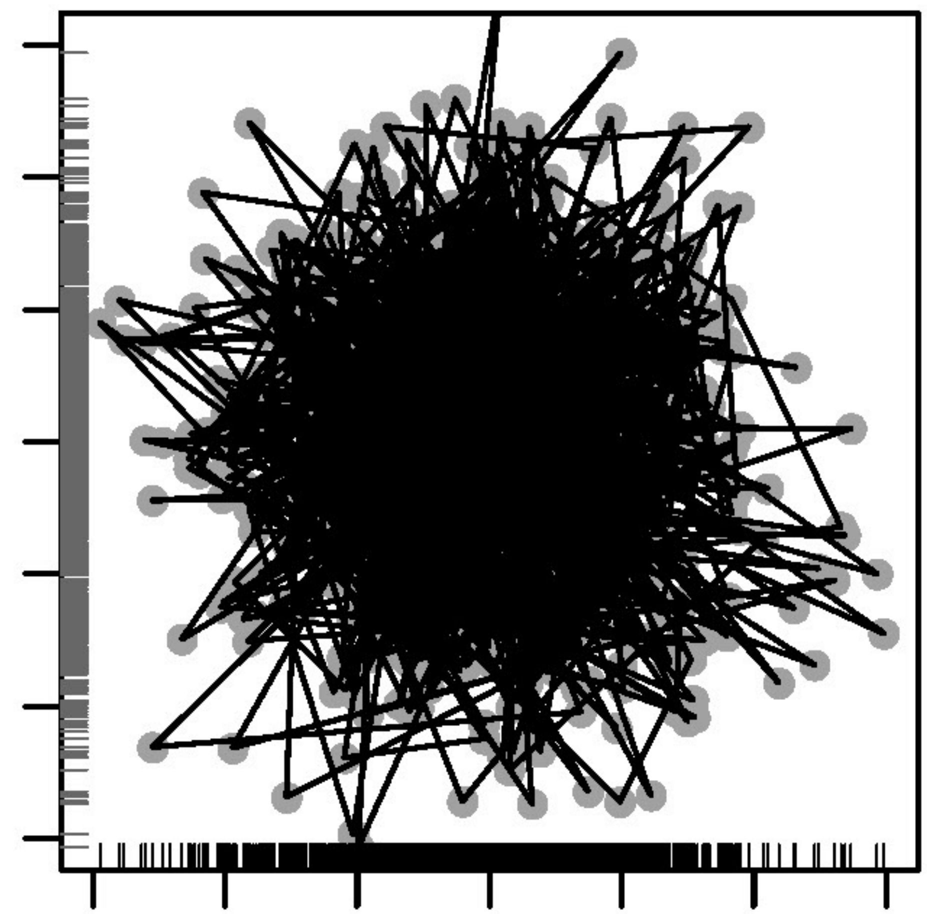
Homogeneous v. non-homogeneous

MINIME SYSTEM

MINIME SYSTEM

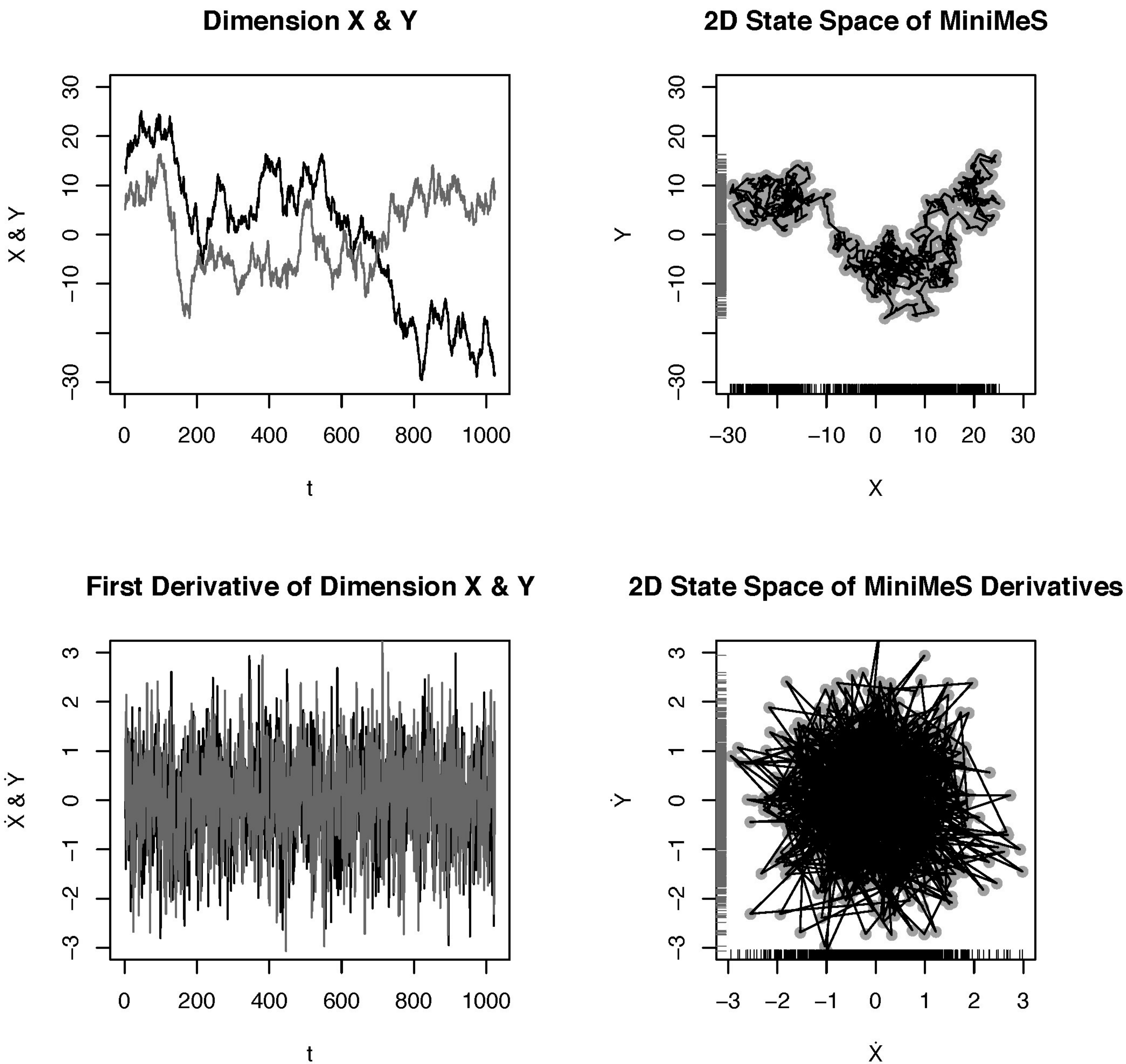


- State = X,Y coordinate
- Minimal Memory System can move around within the boundary.
- When would you infer randomness, when a deterministic rule?
- What kind of succession of states?
- What kind of trajectory through space?

MINIME SYSTEM

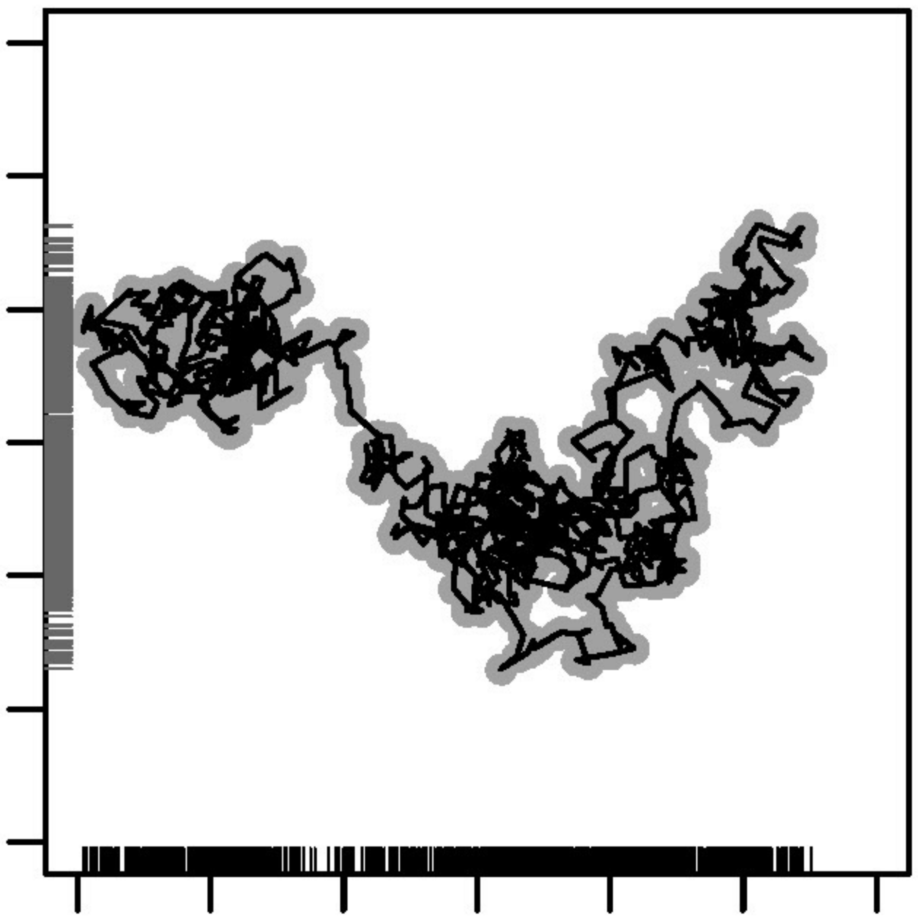
f.hasselman@bsi.ru.nl

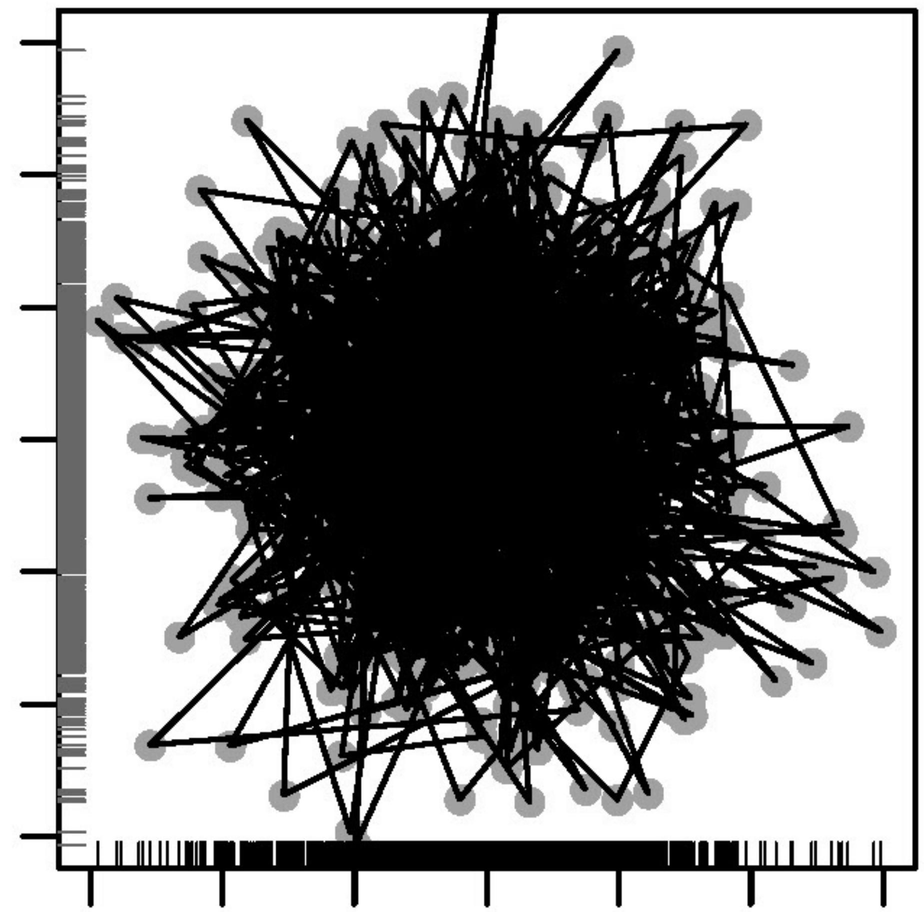


- State Space (X & Y): The degrees of freedom MiniMe has to generate its behaviour (move)
- This is a random walk, Brownian motion: Add a random number drawn from normal distribution to current number.
- Where does the apparent order come from? It's a random process!!!!

‘Simple’ rule reduces degrees of freedom to move around:

Matter has to occupy finite space & movement takes time (no teleportation yet)





Minimal form of ‘physical memory’ through ‘natural computation’: summation / counting

Emergence of structure / temporal correlations / redundancies / dependencies

Brownian motion / Levy flights are very common in nature (diffusion, percolation, foraging)

How to characterise the nature of the dependencies?

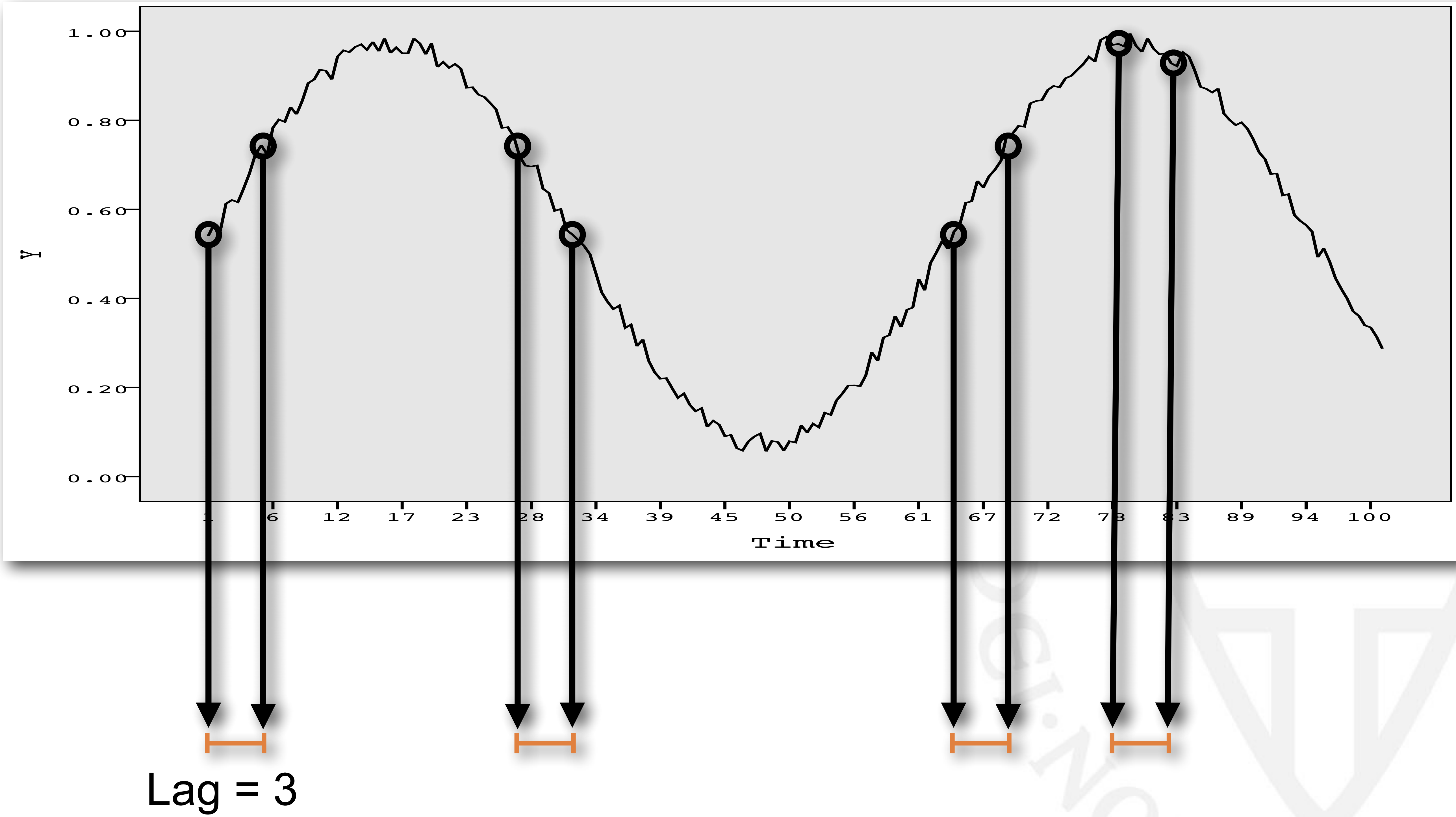
(Partial) Autocorrelation Function - (P)ACF

$$R(k) = \frac{E[(X_i - \mu)(X_{i+k} - \mu)]}{\sigma^2}$$

The average correlation **r** between data points that are a distance (lag) **k** apart in time

This holds only for *stationary*, *random processes.* So X measured here is a *random variable*.

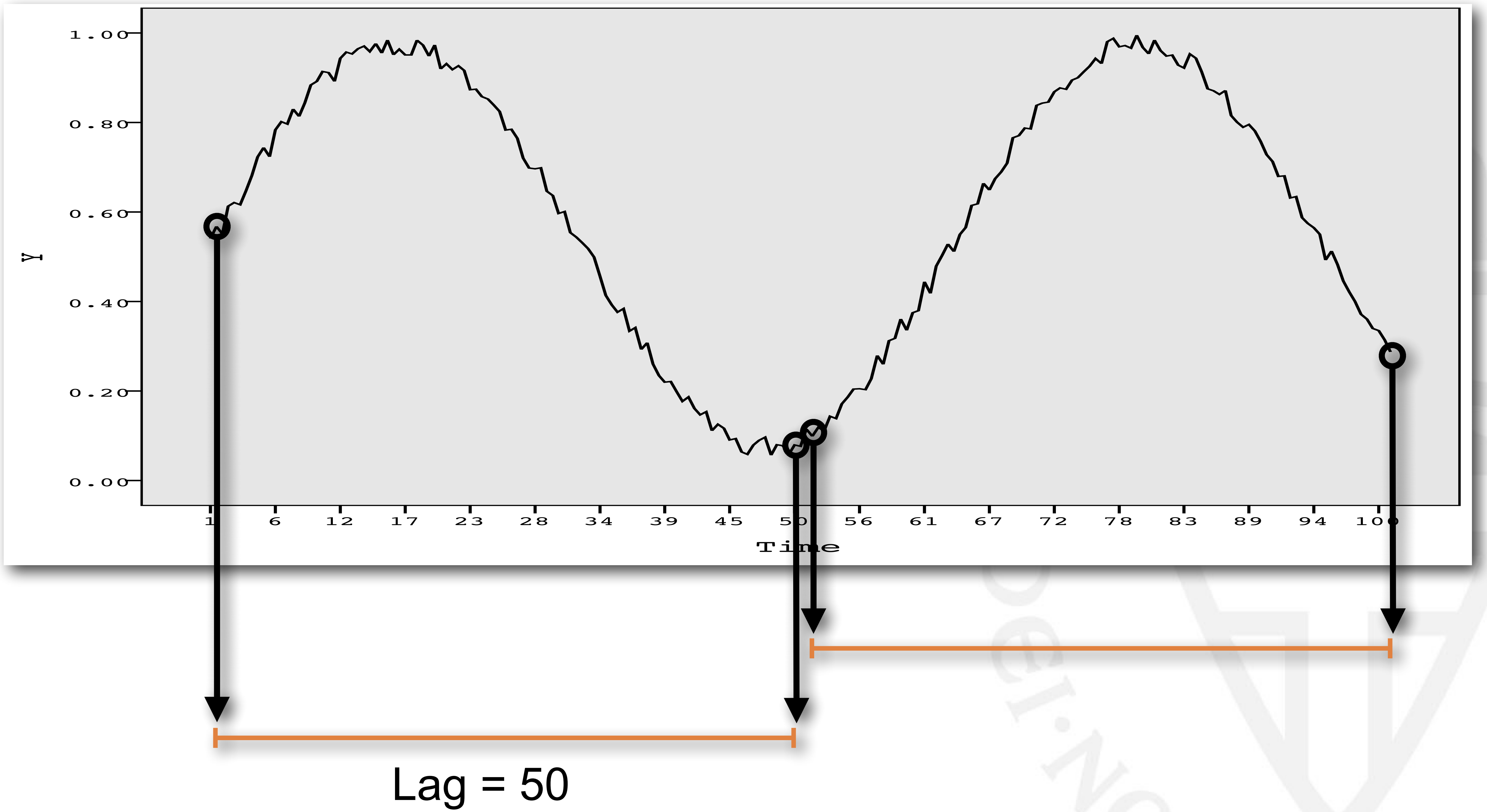
ACF and the Partial ACF are used to decide which AR(fI)MA model you need (how many AR and/or MA parameters you need).



Low or High at lag 3?

r_3= 0.895 (*SD = 0.095*)

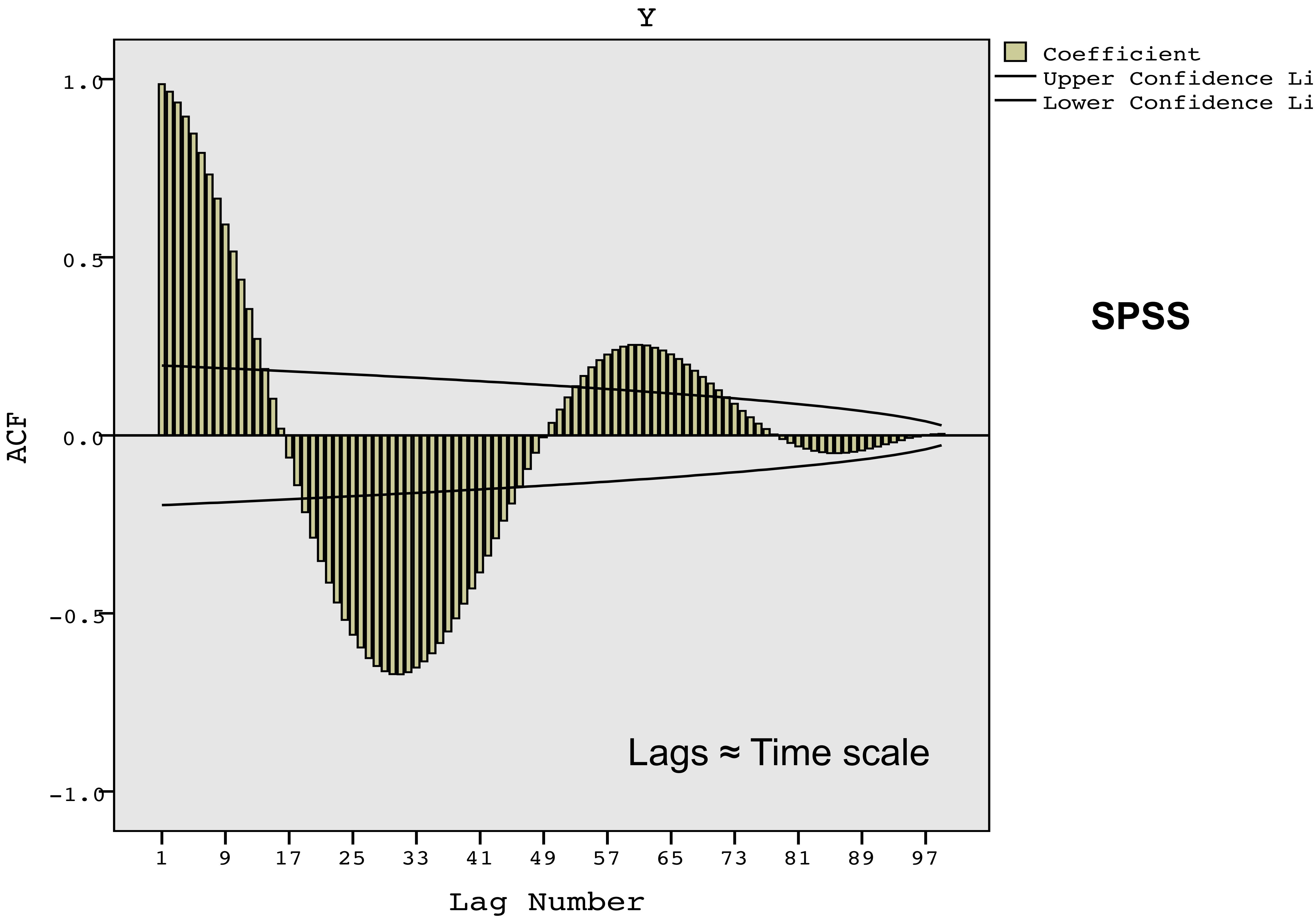
TS length = 100 data points



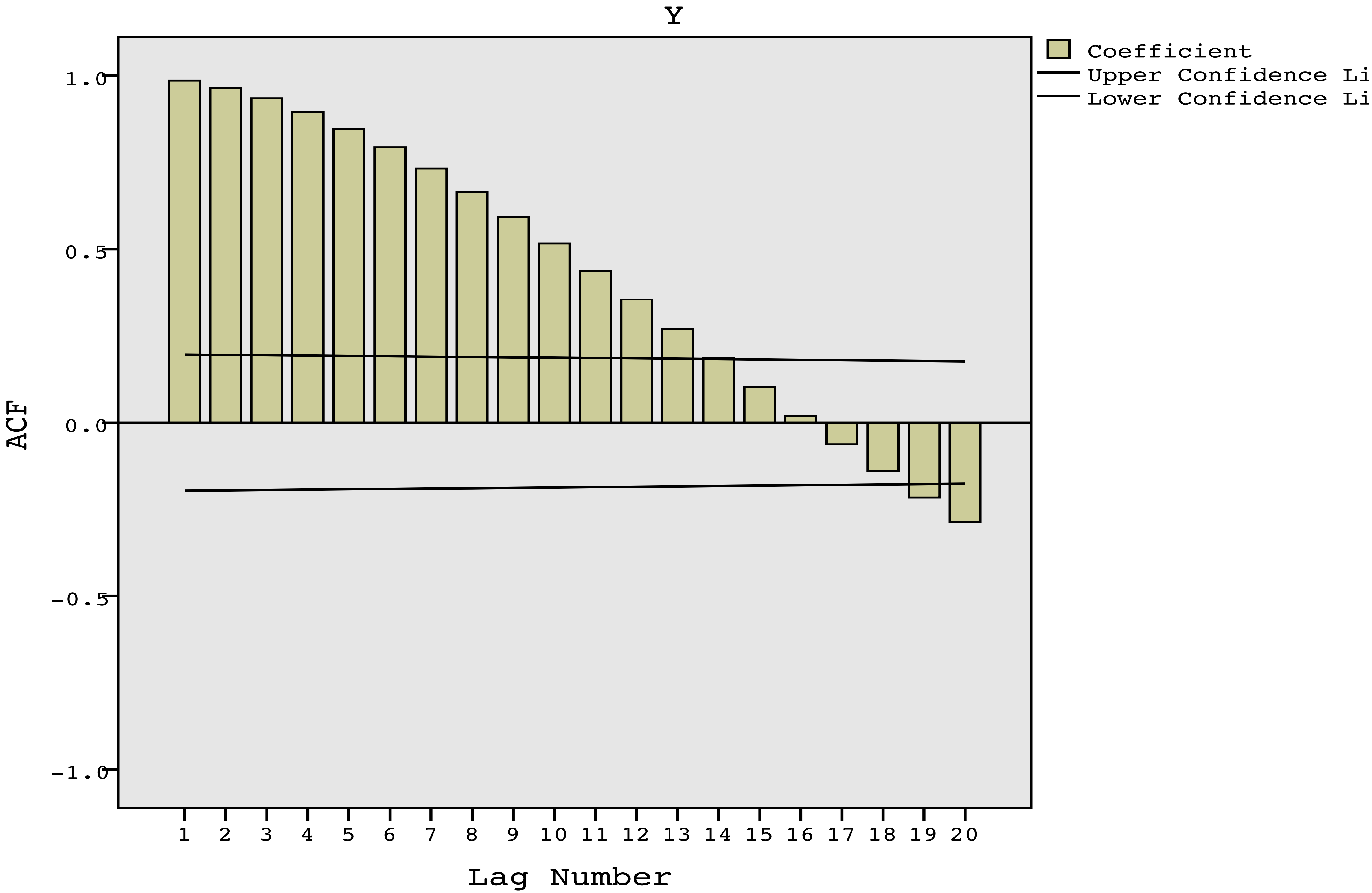
Low or High at lag 50?

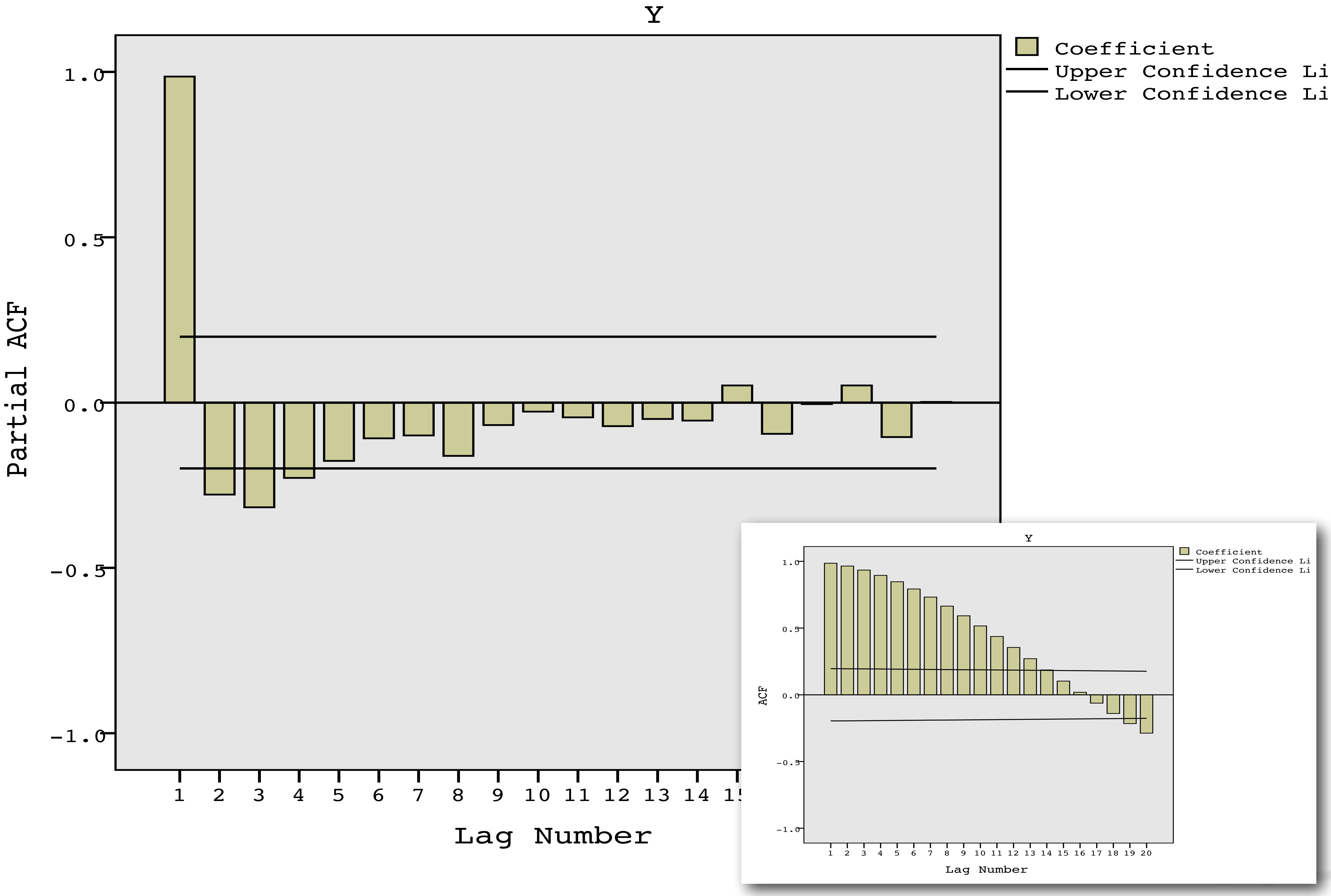
r_{50} = 0.035 (*SD = 0.070*)

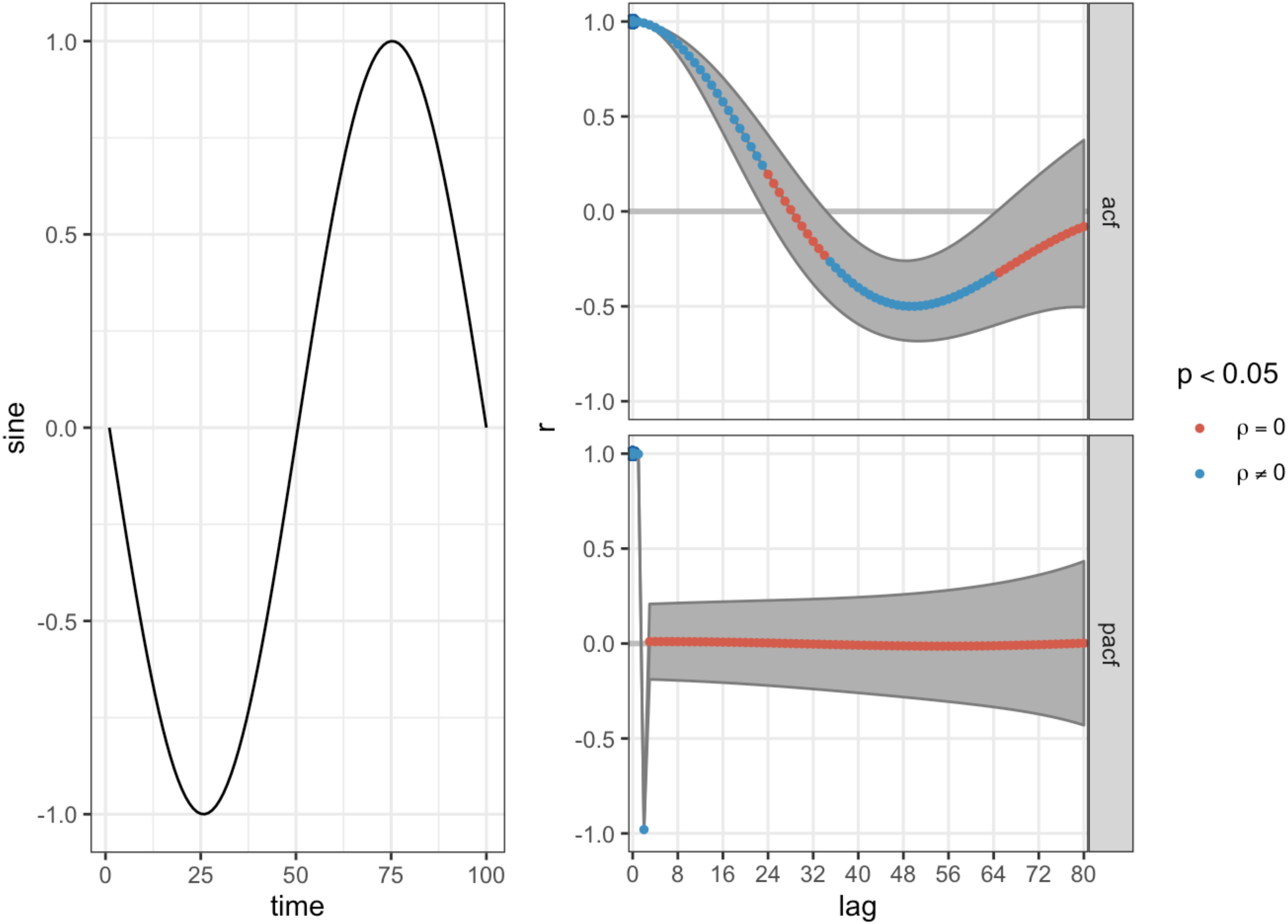
TS length = 100 data points


Y
Coefficient
Upper Confidence Li
Lower Confidence Li
1.0
0.5
0.0
-0.5
-1.0
ACF
1
9
17
25
33
41
49
57
65
73
81
89
97
Lag Number
Lags ≈ Time scale

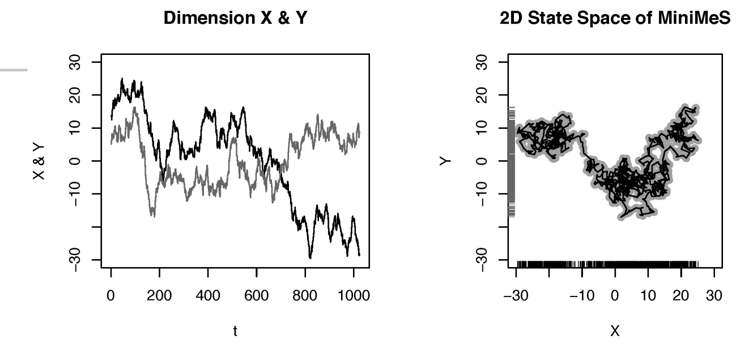
SPSS

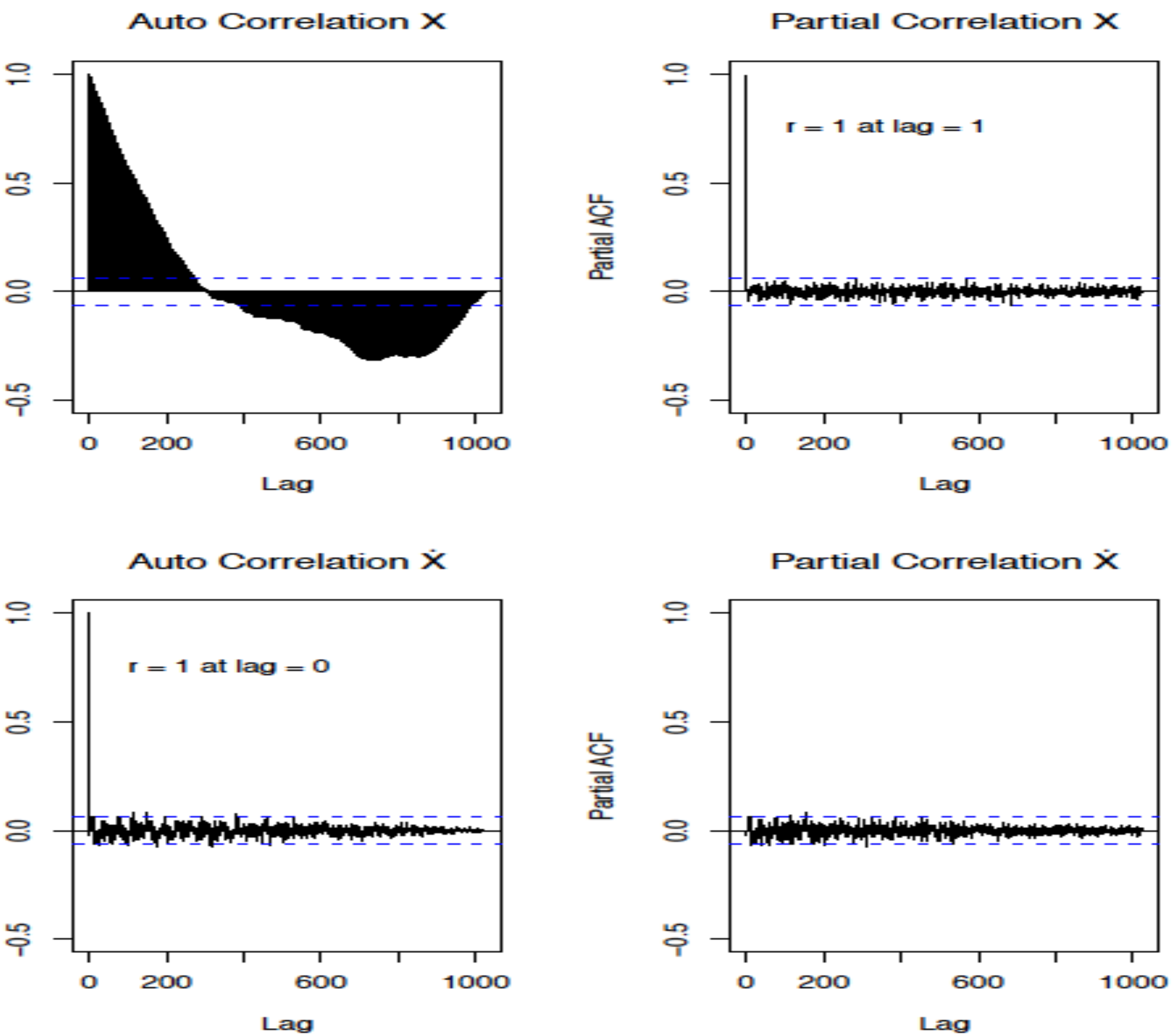

Y
Coefficient
Upper Confidence Li
Lower Confidence Li
ACF
1.0
0.5
0.0
-0.5
-1.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Lag Number


Y
Coefficient
Upper Confidence Li
Lower Confidence Li
1.0
0.5
0.0
-0.5
-1.0
Partial ACF
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Lag Number
Y
Coefficient
Upper Confidence Li
Lower Confidence Li
1.0
0.5
0.0
-0.5
-1.0
ACF
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Lag Number



sine
time
r
lag
acf
pacf
p < 0.05
ρ = 0
ρ ≠ 0





- ACF and PACF: representation of redundancies in the time series
- ACF: Correlations across multiple time scales?
- PACF: Nope, correlation at lag 1

A stochastic linear process model AR(fI)MA

- Notation: **ARIMA(p,d,q)**, these parameters usually take values of 0-2 indicating none, 1 or 2 components.

- **ARIMA(1,2,1)** = means One AutoRegressive parameter, a (filtered-out!) quadratic trend, and 1 Moving Average parameter.

- **AR-part:** The model tries to predict each data point **X** at time **t** based on a constant or intercept (ξ) + a linear combination of previous observations ($\phi_{1...p}$) + random error called random shock (ε):

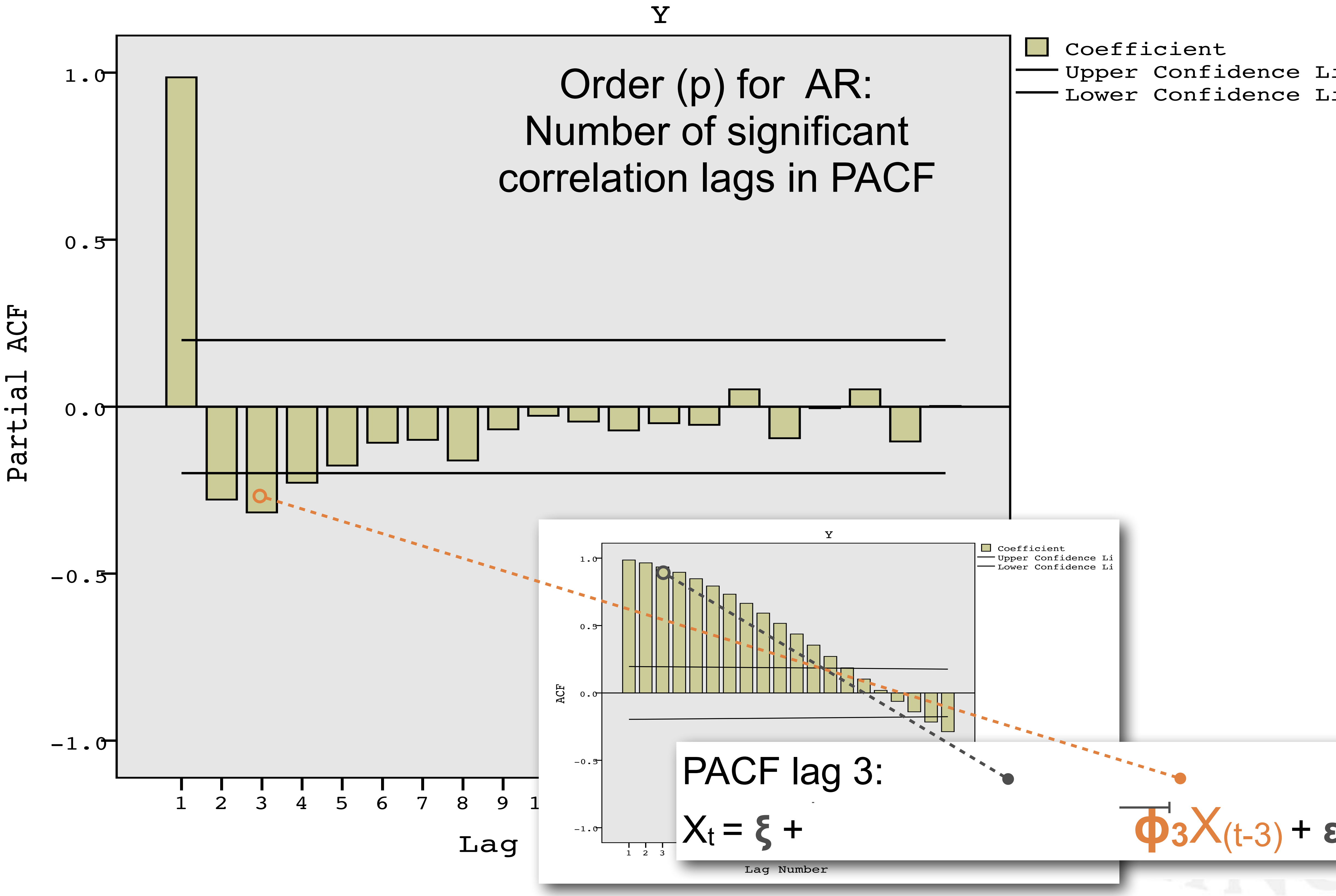
$$X_t = \xi + \phi_1 X_{(t-1)} + \phi_2 X_{(t-2)} + \phi_3 X_{(t-3)} + \ldots + \phi_p X_{(t-p)} + \varepsilon$$

- **MA-part:** The model tries to predict each data point **X** at time **t** based on the average of the time series (μ) + the current random error (ε_t) - a linear combination of previously observed random error ($\theta_{1...q}$):

$$X_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{(t-1)} - \theta_2 \varepsilon_{(t-2)} - \theta_3 \varepsilon_{(t-3)} - \ldots - \theta_p \varepsilon_{(t-q)}$$

- **I(ntegration)-part:** The data must be stationary (and more!): 1) constant mean, 2) constant variance, 3) constant autocorrelation *through* time. This part removes correlations between data points that are a particular distance (lag) apart by differencing the time series… in other words: Trends in the data are filtered out.

- How many parameters do you need? … ACF and PACF



Order (p) for AR:
Number of significant correlation lags in PACF

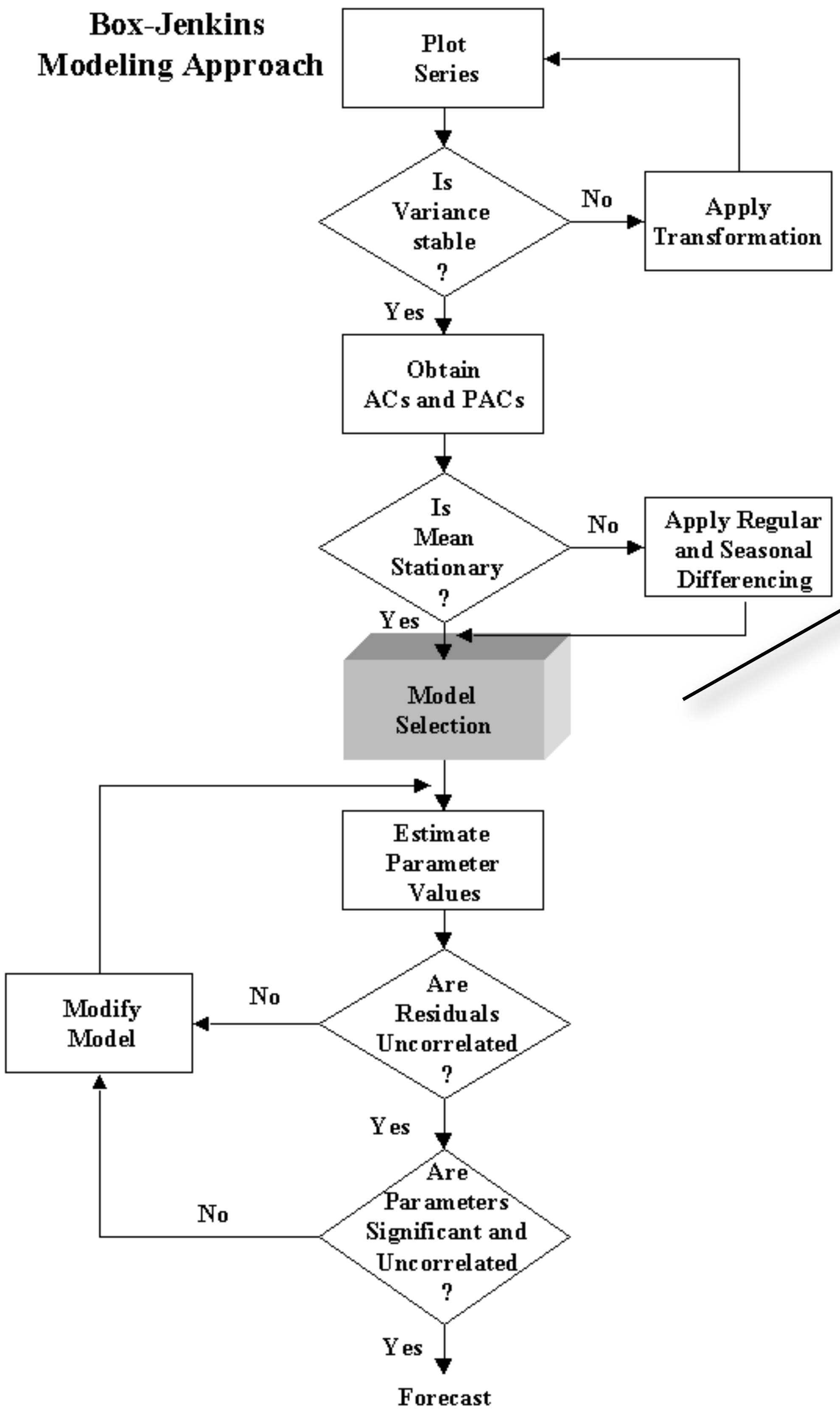
PACF lag 3:

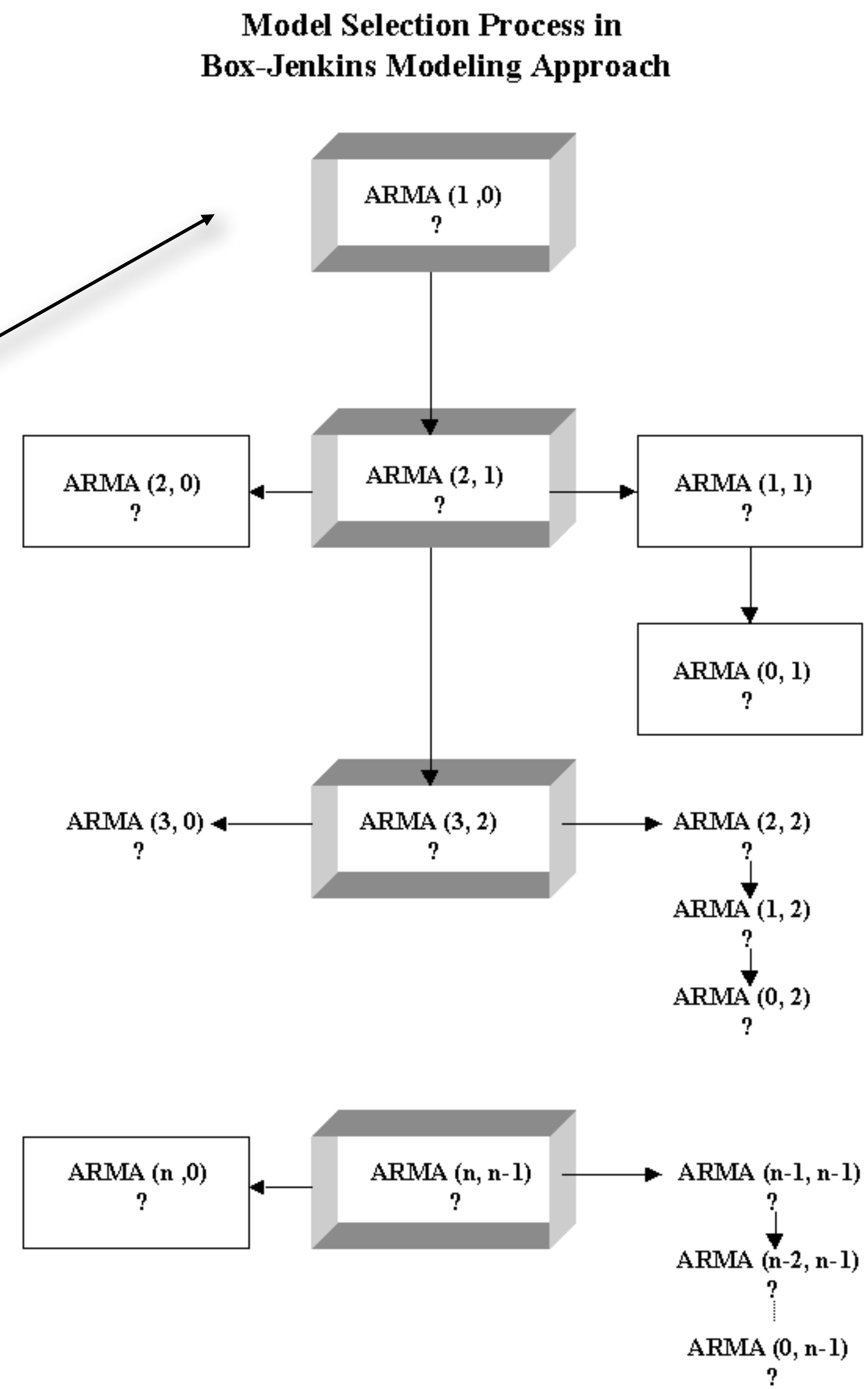
$X_t = \xi + \phi_3 X_{(t-3)} + \varepsilon$

Behavioural Science Institute
Radboud University Nijmegen

Rules of thumb
(http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc446.htm)

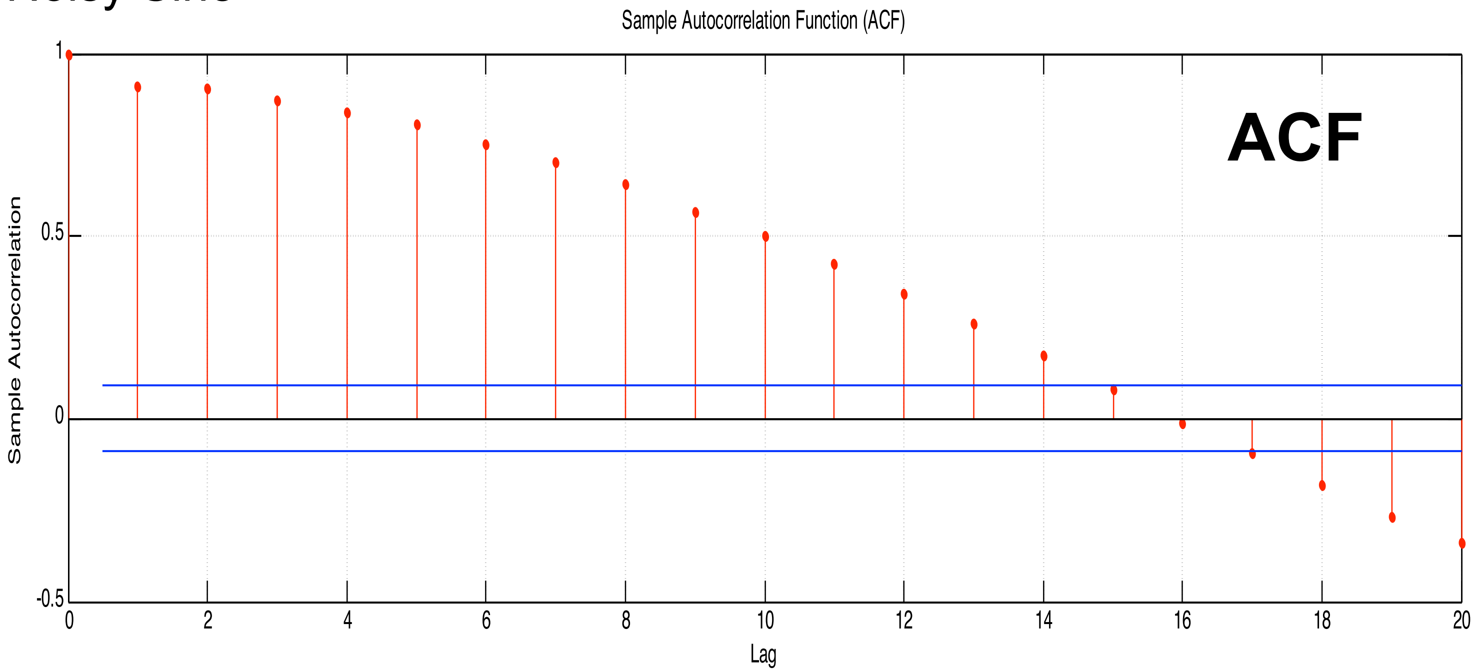
SHAPE of ACF	INDICATED MODEL
Exponential, decaying to zero	Autoregressive model. Use the partial autocorrelation plot to identify the order of the autoregressive model.
Alternating positive and negative, decaying to zero	Autoregressive model. Use the partial autocorrelation plot to help identify the order.
One or more spikes, rest are essentially zero	Moving average model, order identified by where plot becomes zero.
Decay, starting after a few lags	Mixed autoregressive and moving average model.
All zero or close to zero	Data is essentially random.
High values at fixed intervals	Include seasonal autoregressive term.
No decay to zero	Series is not stationary.

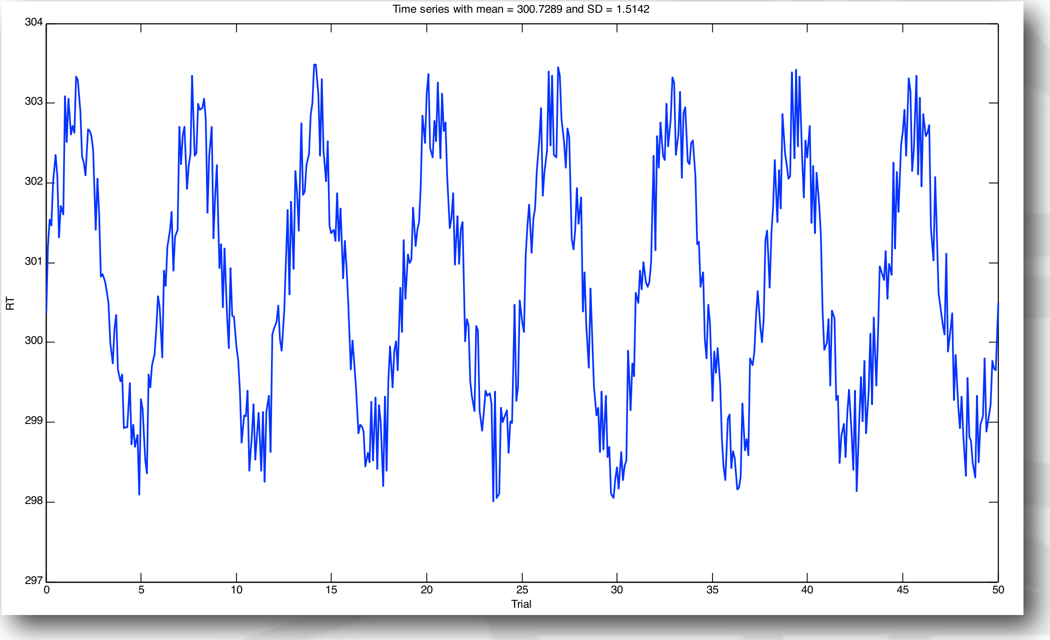

Box-Jenkins
Modeling Approach
Plot Series
Is Variance stable ?
No
Apply Transformation
Yes
Obtain ACs and PACs
Is Mean Stationary ?
No
Apply Regular and Seasonal Differencing
Yes
Model Selection
Estimate Parameter Values
Are Residuals Uncorrelated ?
No
Modify Model
Yes
Are Parameters Significant and Uncorrelated ?
No
Yes
Forecast


Model Selection Process in
Box-Jenkins Modeling Approach
ARMA (1 ,0) ?
ARMA (2, 0) ?
ARMA (2, 1) ?
ARMA (1, 1) ?
ARMA (0, 1) ?
ARMA (3, 0) ?
ARMA (3, 2) ?
ARMA (2, 2) ?
ARMA (1, 2) ?
ARMA (0, 2) ?
ARMA (n ,0) ?
ARMA (n, n-1) ?
ARMA (n-1, n-1) ?
ARMA (n-2, n-1) ?
ARMA (0, n-1) ?

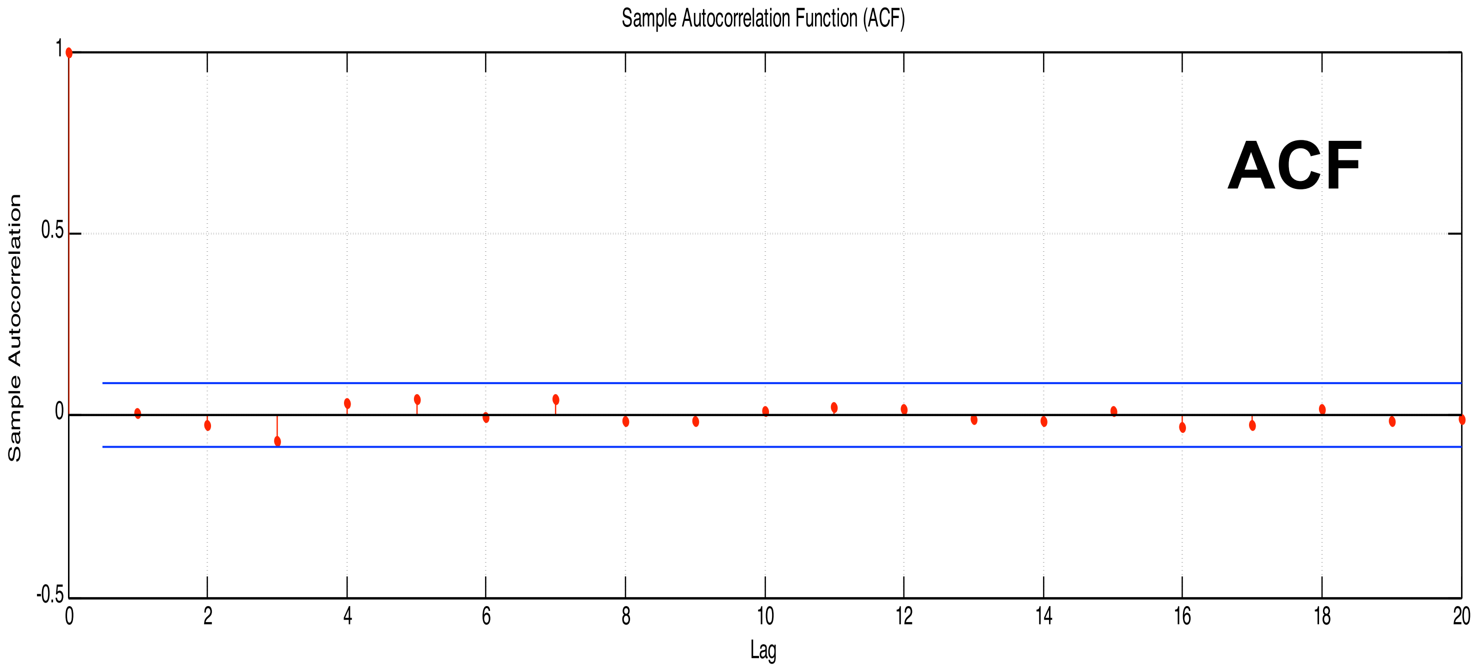
Randomising temporal order = Destroying correlations in the data

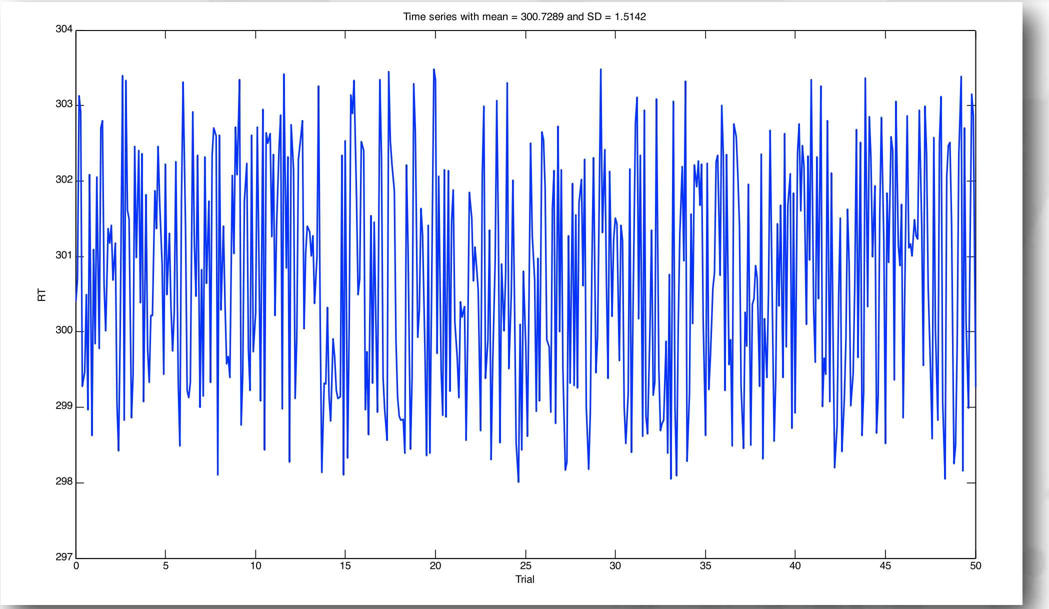
Noisy Sine





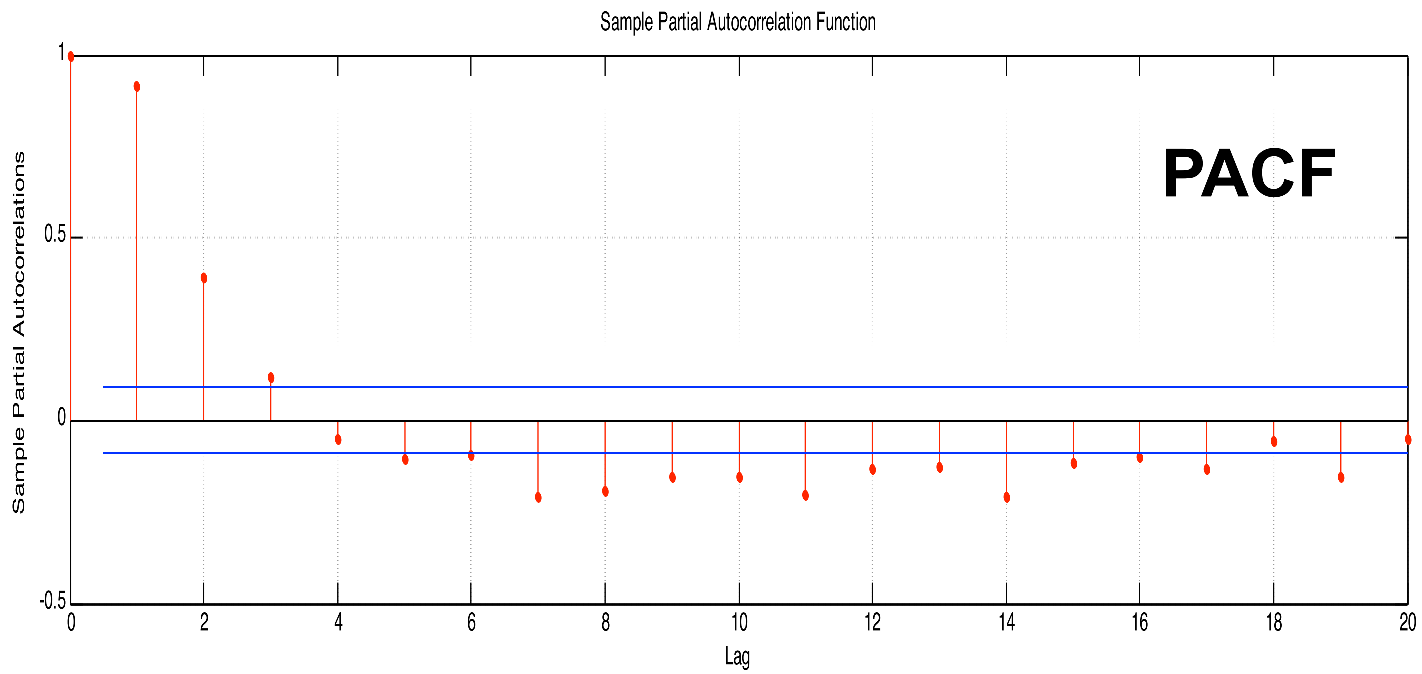
Randomised Noisy Sine

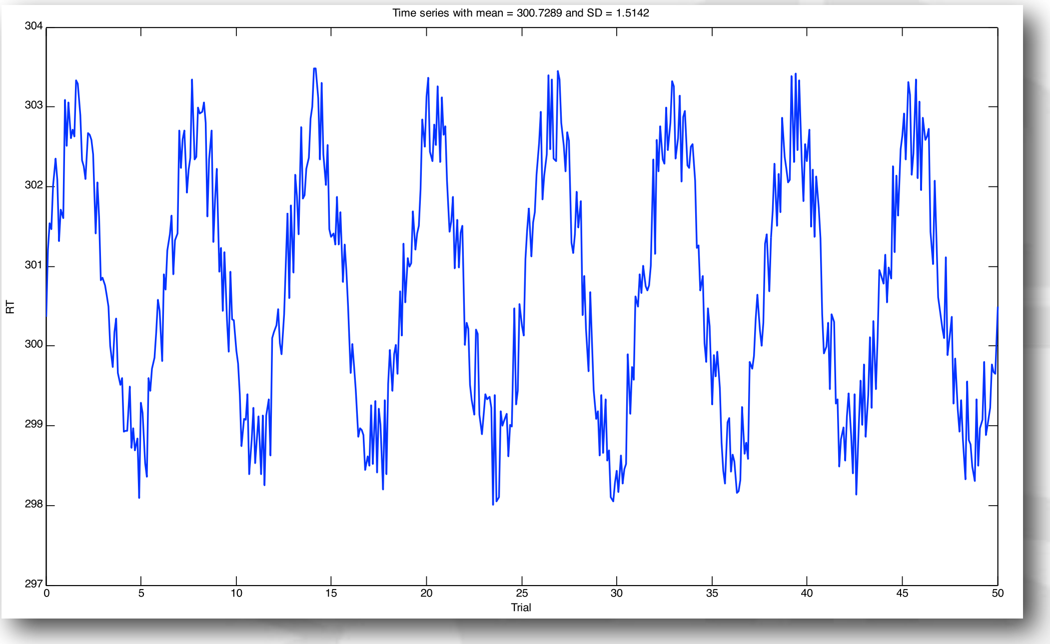




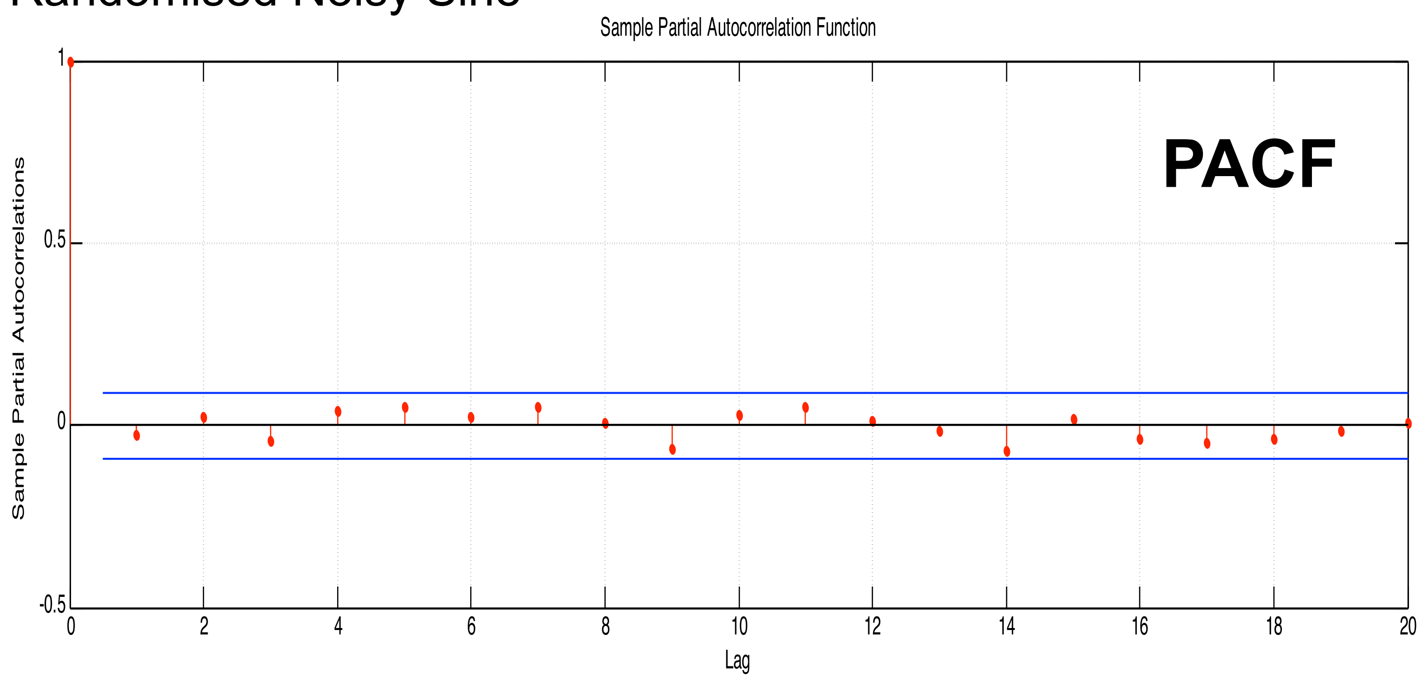
Randomising temporal order = Destroying correlations in the data

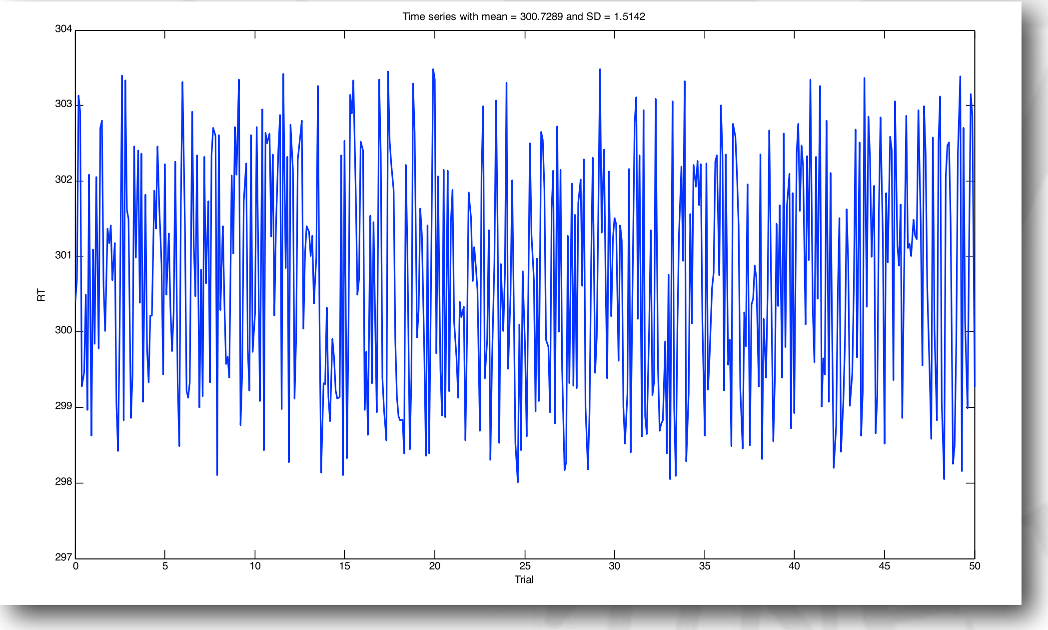
Noisy Sine





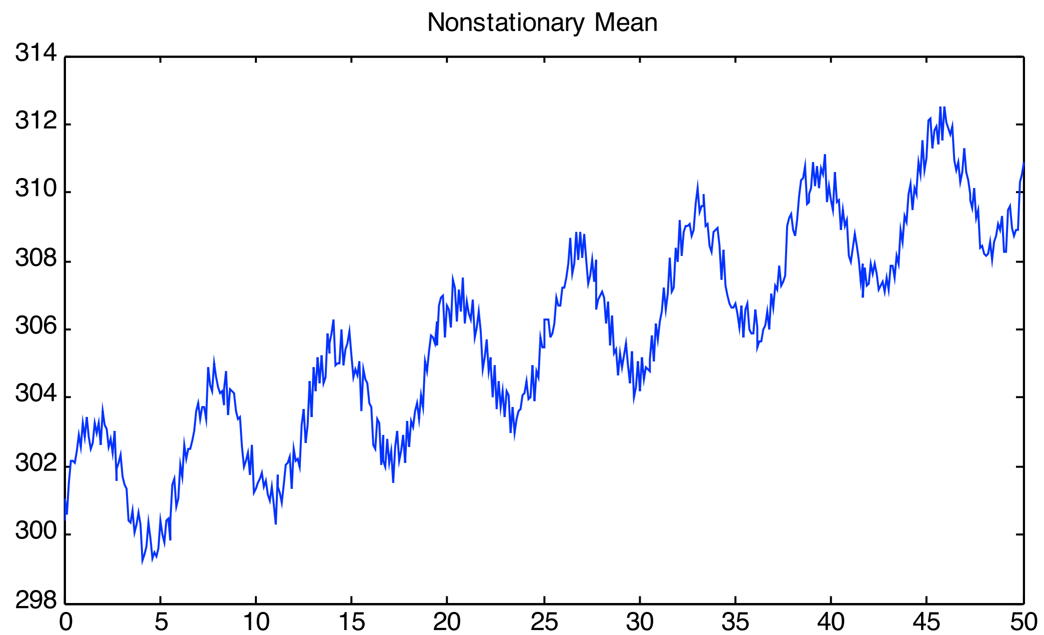
Randomised Noisy Sine

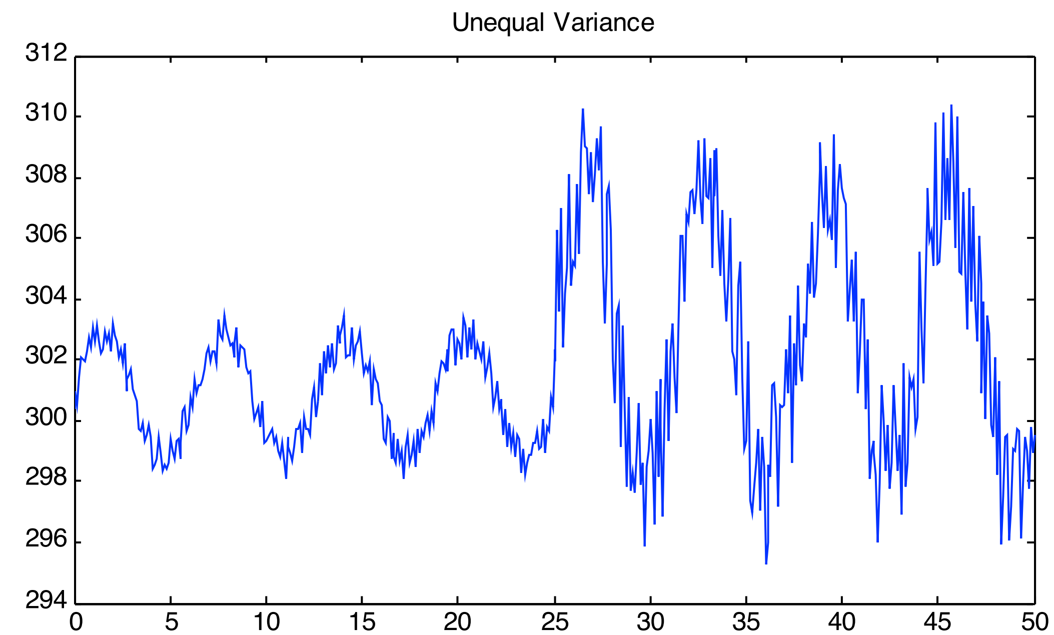


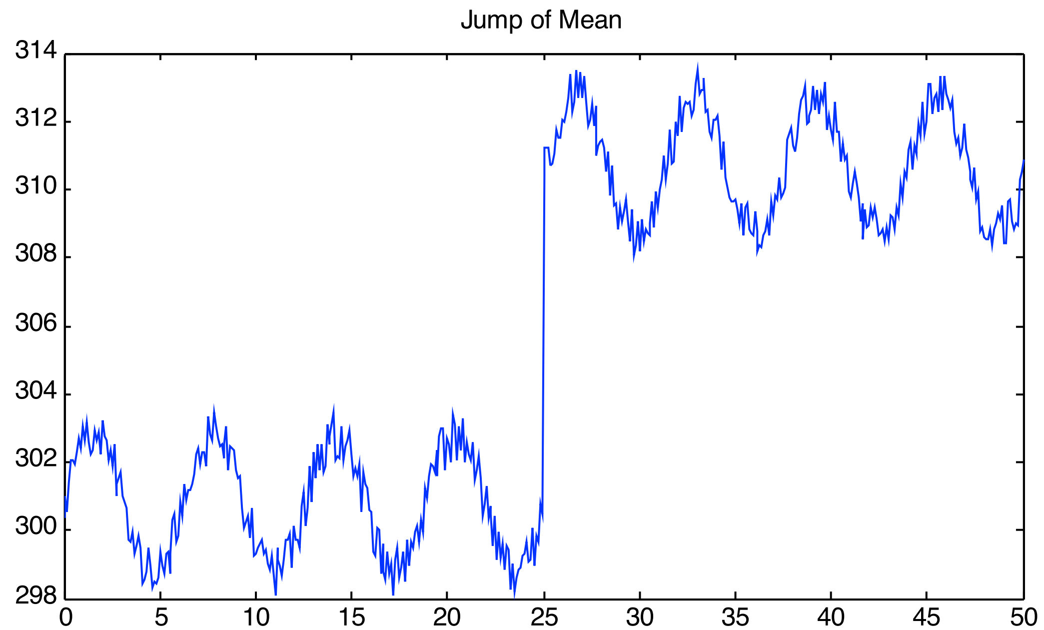


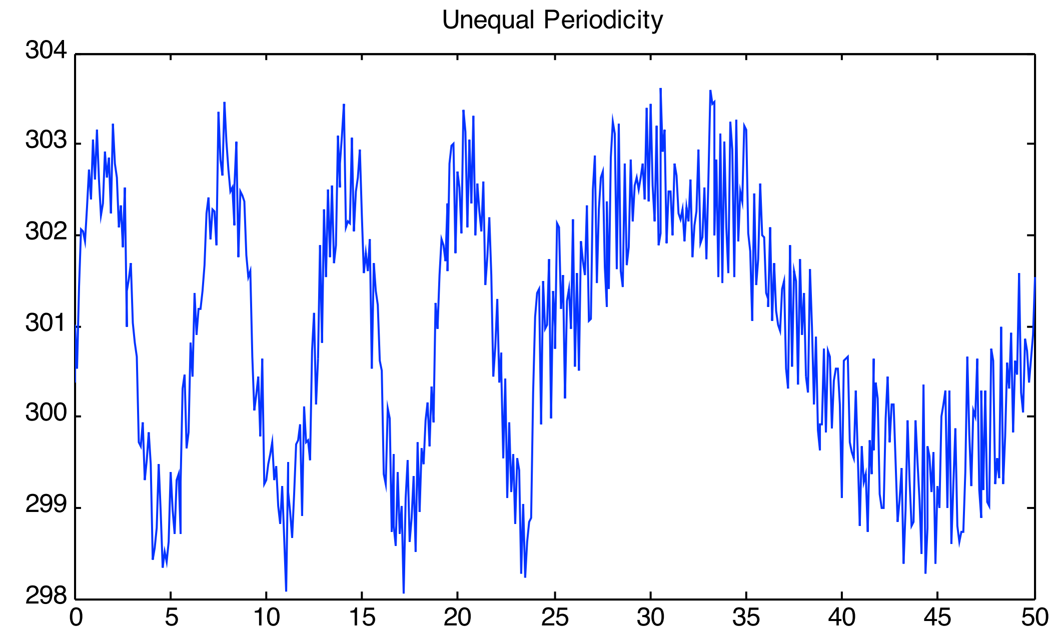
Behavioural Science Institute
Radboud University Nijmegen

Problems with ARfIMA (data assumptions)









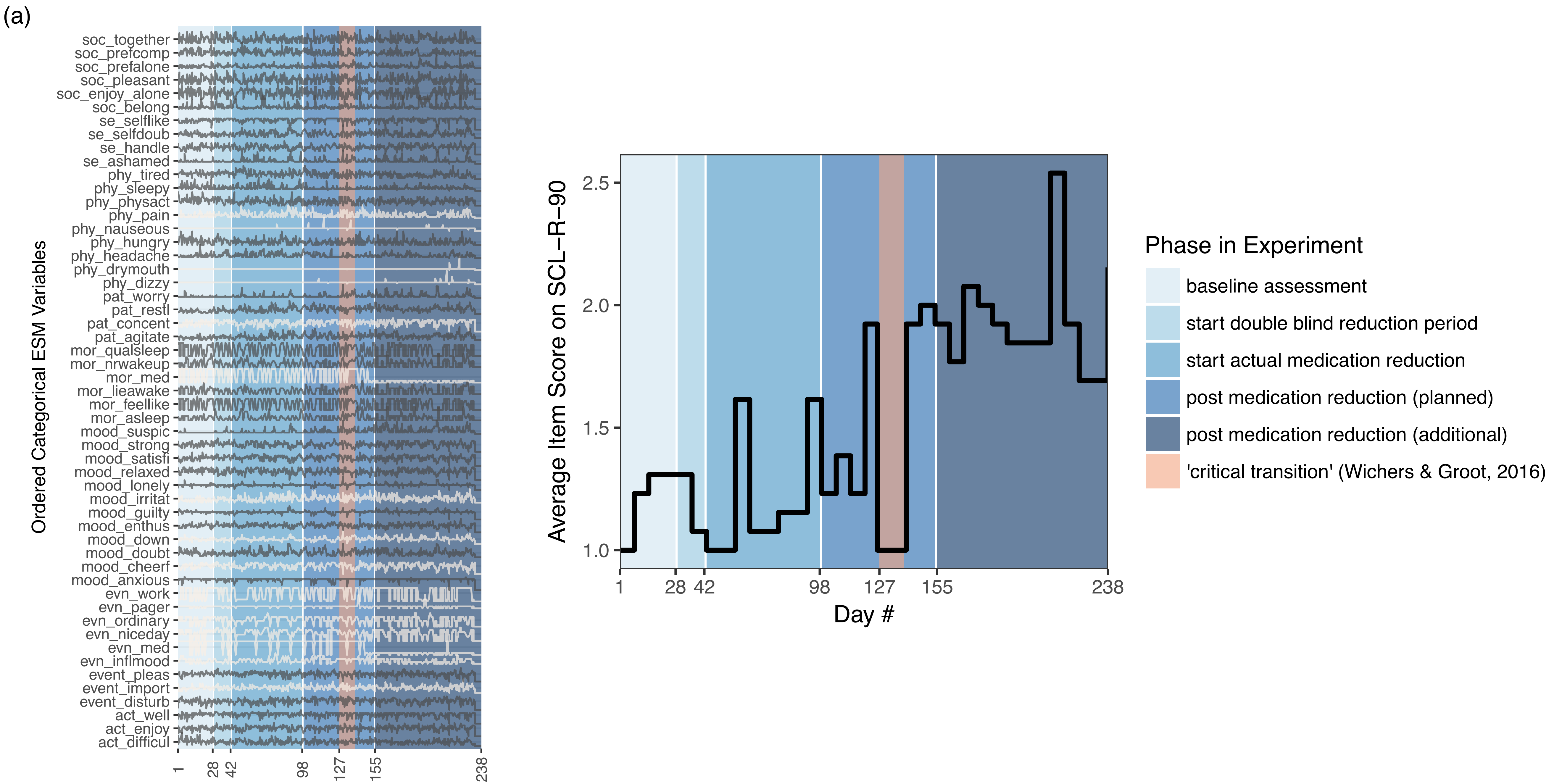
Testing for ergodicity

Testing for **stationarity**

Testing for **homogeneity**

http://fredhasselman.com/post/2017-05-19-testing-assumptions-of-the-data-generating-process-underlying-experience-sampling/

“Critical Slowing Down as a Personalized Early Warning Signal for Depression”



Wichers, M., Groot, P. C., Psychosystems, ESM Grp, & EWS Grp (2016). Critical Slowing Down as a Personalized Early Warning Signal for Depression. Psychotherapy and psychosomatics, 85(2), 114-116. DOI: 10.1159/000441458

Kossakowski, J., Groot, P., Haslbeck, J., Borsboom, D., and Wichers, M. (2017). Data from ‘critical slowing down as a personalized early warning signal for depression’. Journal of Open Psychology Data, 5(1).

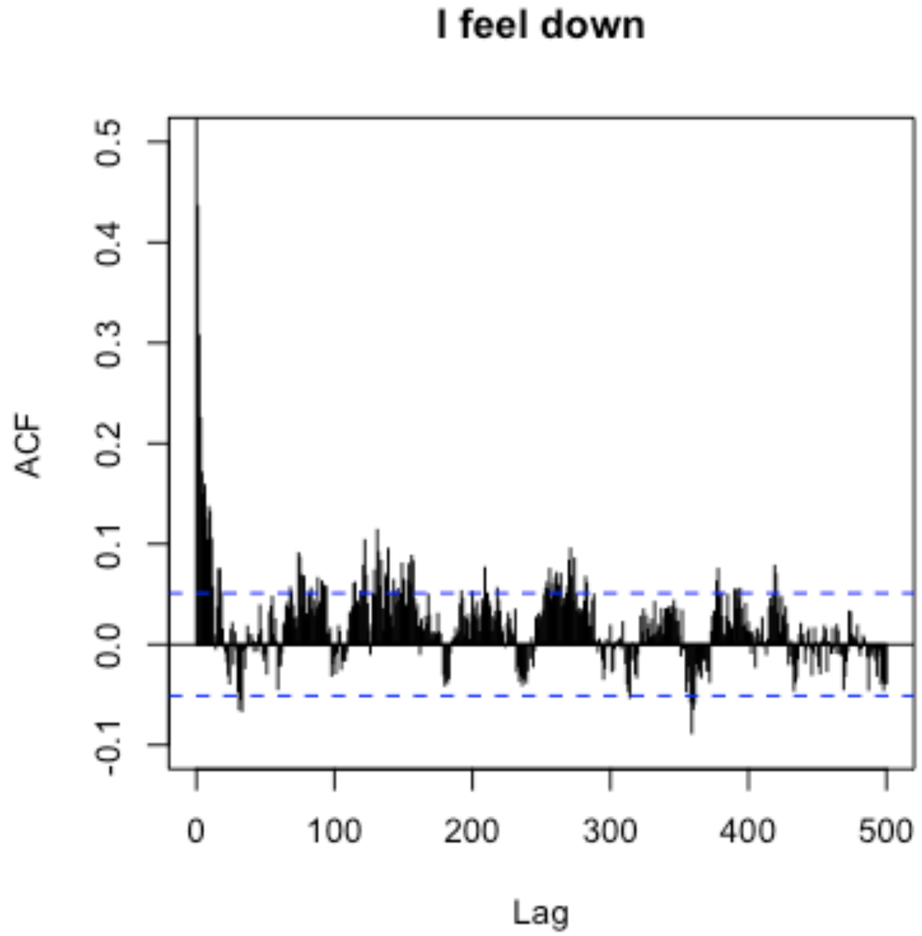
Rank Version of von Neumann's Ratio Test for Randomness | **Kwiatkowski–Phillips–Schmidt–Shin (KPSS)**

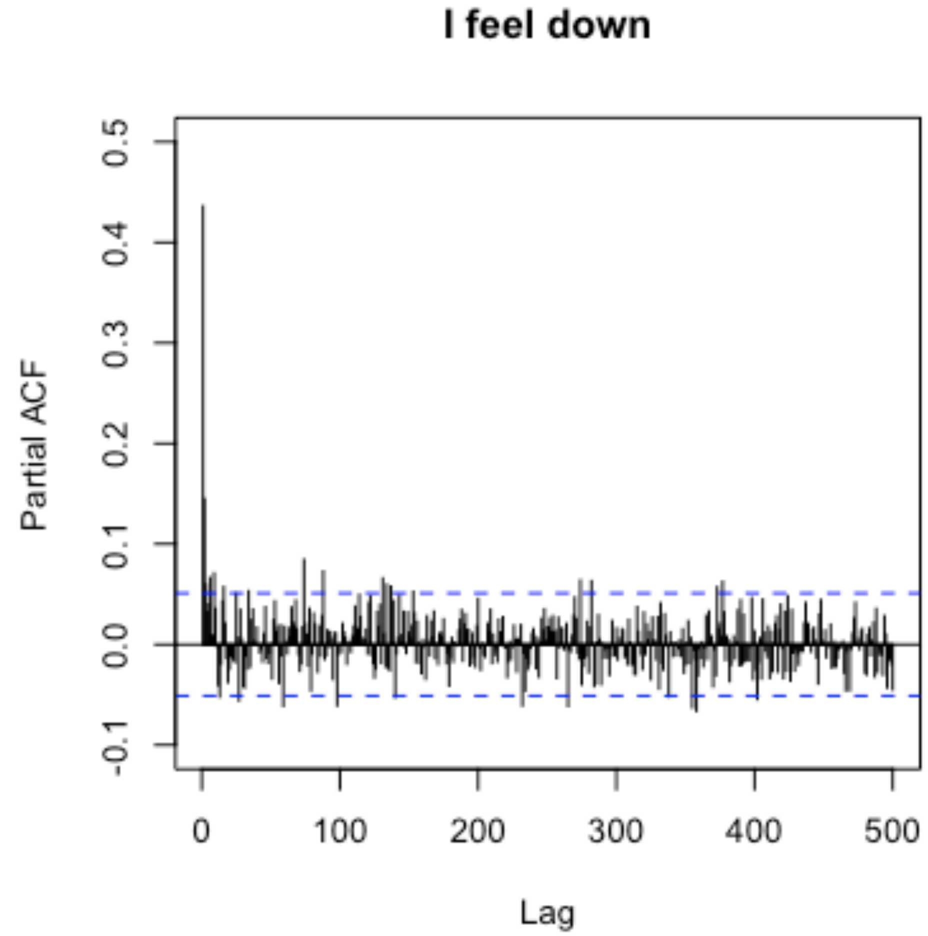
	Bartels rank test H0 = Random H1 = Non-random		**KPSS test H0 = Level Stationary H1 = Unit root**		**KPSS test H0 = Trend Stationary H1 = Unit root**		**Significant partial autocorrelations**	
Item	*All data*	*Subset*	*All data*	*Subset*	*All data*	*Subset*	*Lag 2-99*	*Lag 100-1000*
I feel relaxed	<.001*	<.001*	0.092	0.046	0.036	0.021	2	6
I feel down	<.001*	<.001*	<.010*	0.100	0.100	0.100	8	8
I feel irritated	<.001*	<.001*	<.010*	0.052	<.010*	0.100	5	7
I feel satisfied	<.001*	<.001*	0.100	0.019	0.100	0.098	2	4
I feel lonely	<.001*	<.001*	<.010*	0.100	0.100	0.100	5	9
I feel anxious	<.001*	<.001*	<.010*	0.100	0.100	0.100	8	11
I feel enthusiastic	<.001*	<.001*	0.100	0.100	0.100	0.100	4	6
I feel suspicious	<.001*	<.001*	<.010*	0.061	0.041	0.027	9	9
I feel cheerful	<.001*	<.001*	0.100	0.059	0.100	0.046	4	6
I feel guilty	<.001*	<.001*	<.010*	<.010*	0.094	0.100	7	7
I feel indecisive	<.001*	<.001*	0.100	<.010*	0.050	0.100	7	7
I feel strong	<.001*	<.001*	0.100	0.021	0.100	0.100	6	6
I feel restless	<.001*	<.001*	<.010*	0.070	<.010*	0.075	11	4
I feel agitated	<.001*	<.001*	<.010*	0.100	<.010*	0.100	6	5
I worry	<.001*	<.001*	<.010*	0.100	0.100	0.100	10	11
I can concentrate well	<.001*	<.001*	<.010*	<.010*	0.100	0.100	4	8
I like myself	<.001*	<.001*	0.100	<.010*	0.082	0.100	5	5
I am ashamed of myself	<.001*	<.001*	<.010*	0.100	0.100	0.100	8	6
I doubt myself	<.001*	<.001*	0.048	0.100	0.093	0.100	7	5
I can handle anything	<.001*	<.001*	0.055	0.047	0.100	0.100	4	8
I am hungry	0.068	0.068	<.010*	0.020	<.010*	0.049	6	2
I am tired	<.001*	<.001*	<.010*	0.100	0.079	0.978	11	5
I am in pain	<.001*	<.001*	0.100	0.024	<.010*	0.100	4	2
I feel dizzy	0.854		<.010*		0.050		6	7
I have a dry mouth	0.958		0.029		0.042		1	8
I feel nauseous	0.854		0.100		0.100		4	9
I have a headache	<.001*	0.8544	0.018	0.020	<.010*	0.100	7	4
I am sleepy	<.001*	0.958	<.010*	0.011	<.010*	0.100	7	4
From the last beep onwards I was physically active	<.001*	0.854	<.010*	0.100	<.010*	0.100	3	3
Sum of significant tests (%)	25 (86%)	22 (85%)	16 (55%)	4 (15%)	8 (28%)	0 (0%)		

Note.

N = 1476 for all data. N = 292 for the subset **[= START ACTUAL REDUCTION]**.

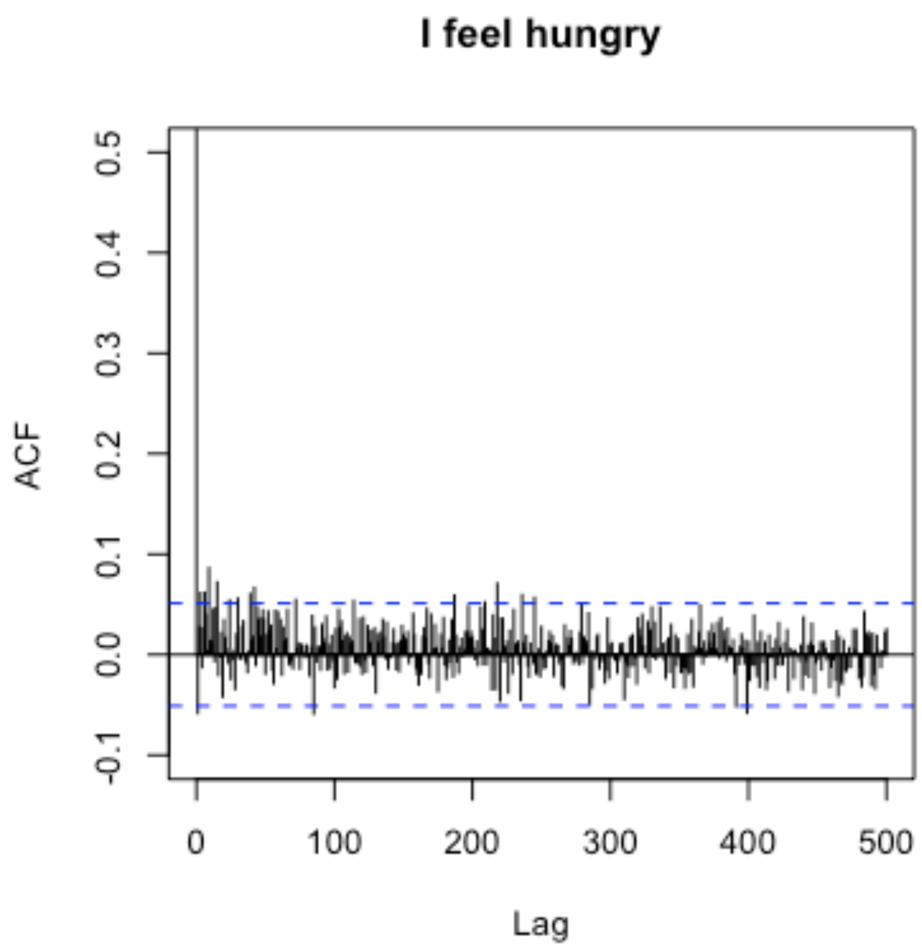
* indicates statistically significant test statistics. For Bartels rank test, results were considered significant for $p<.002$. The KPSS test only provides p-values in between .01 and .10. For the KPSS test, $p<.010$ was considered significant. Three items showed no variance during the baseline period included in the subset and were therefore omitted from analysis of the subset.

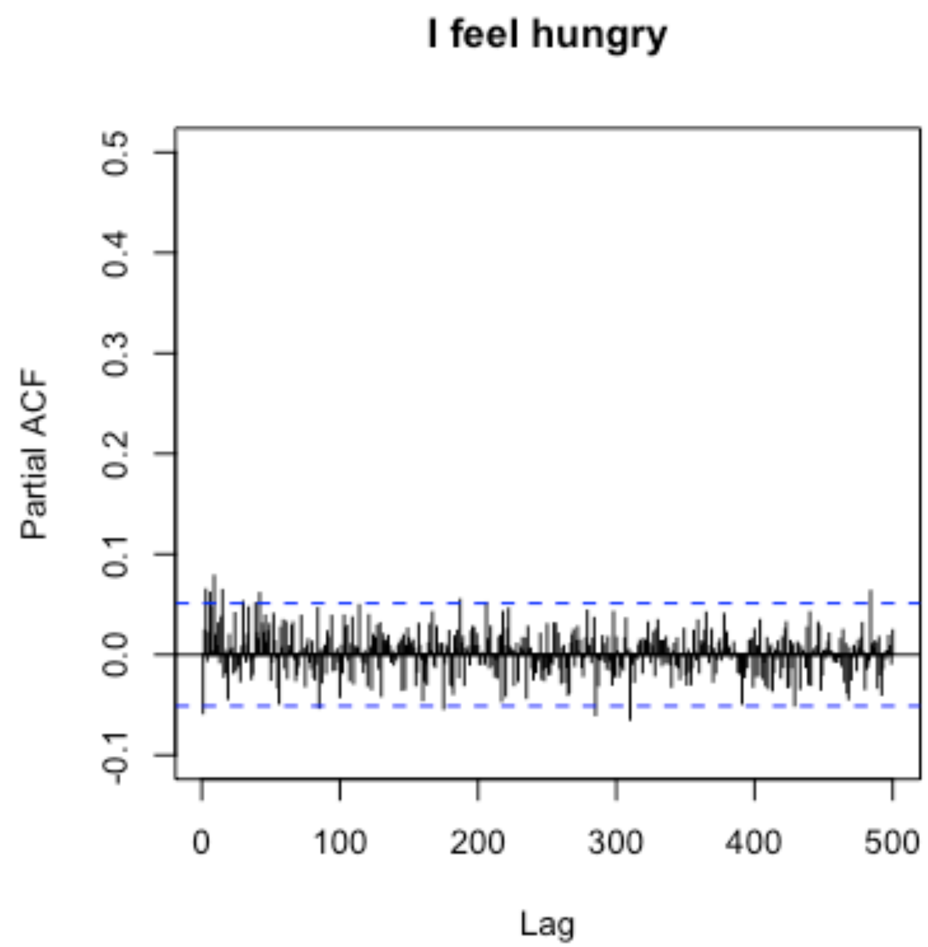




Questions abt. *mental internal states* like ***mood*** resemble non-ergodic processes:

- long memory
- non-stationary
- non-homogeneous
- non-stationary ACF





Questions abt. *physical internal states* like ***hunger*** resemble ergodic processes:

- no long memory
- stationary
- homogeneous
- stationary ACF

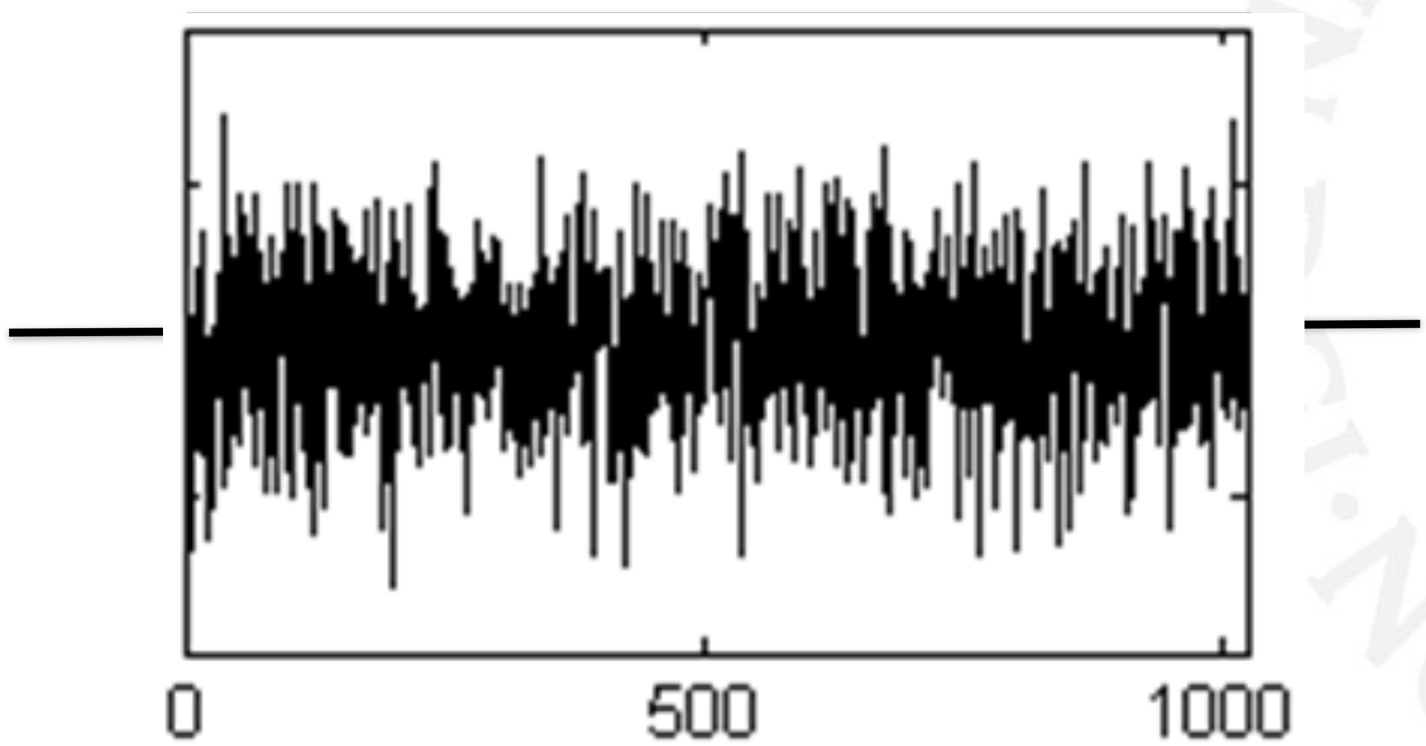
Intuitive Notion of Fractal Dimension

Relative Roughness

Entropy

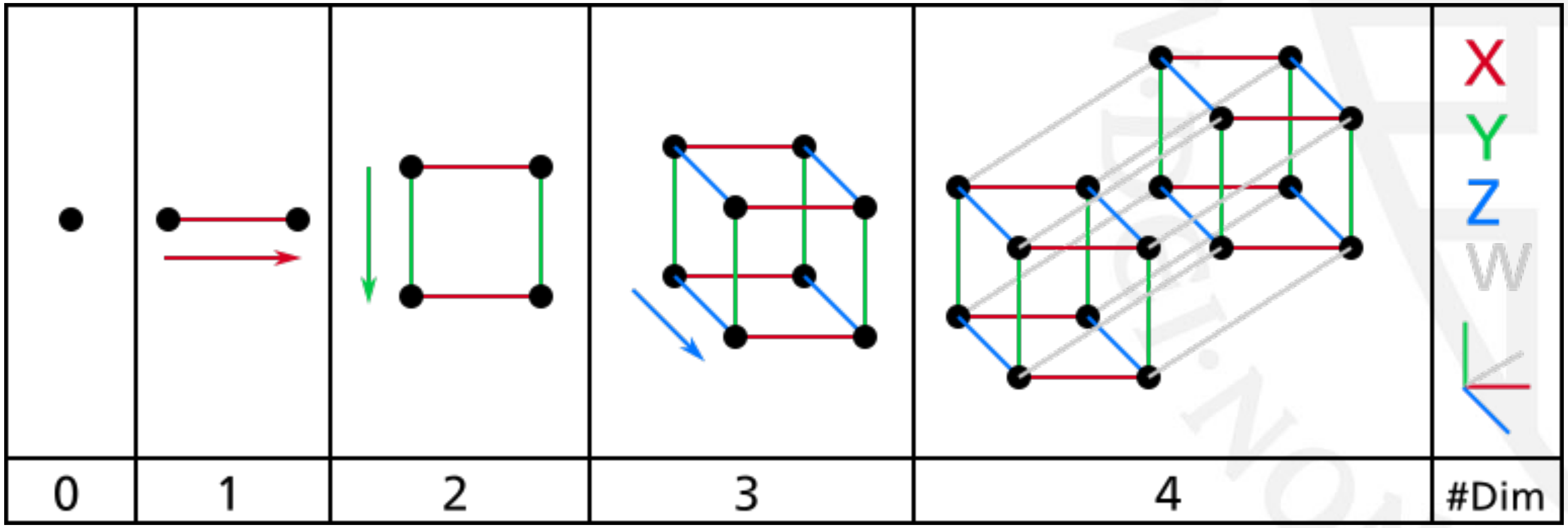
Fractal dimension

What is the dimension of a line?
What is the dimension of a rectangle?
What is the dimension of random noise?

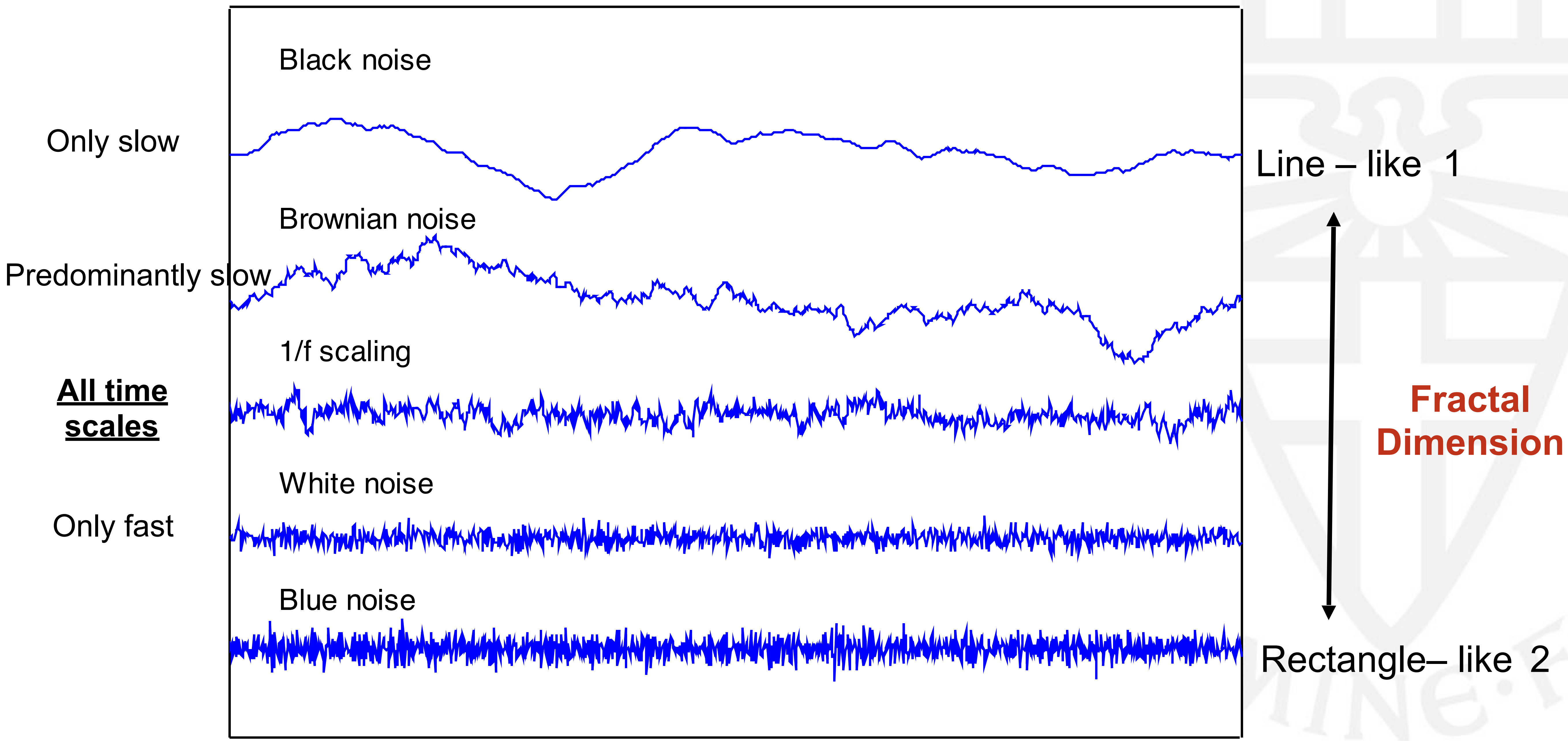


Fractal dimension

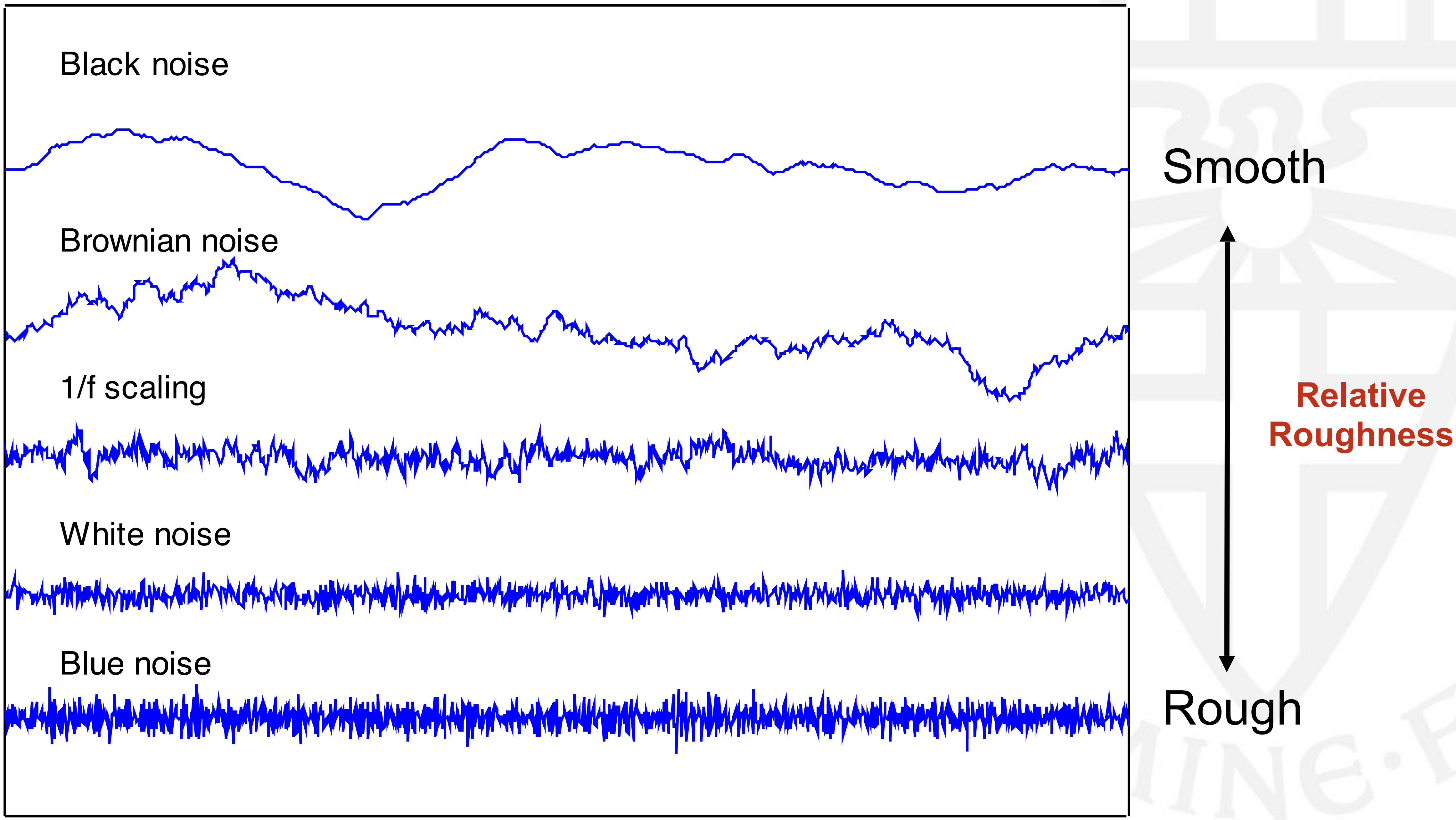
What is the dimension of a line?
What is the dimension of a rectangle?
What is the dimension of random noise?



Temporal properties of variability: Fractal Dimension



Temporal properties of variability: Relative Roughness

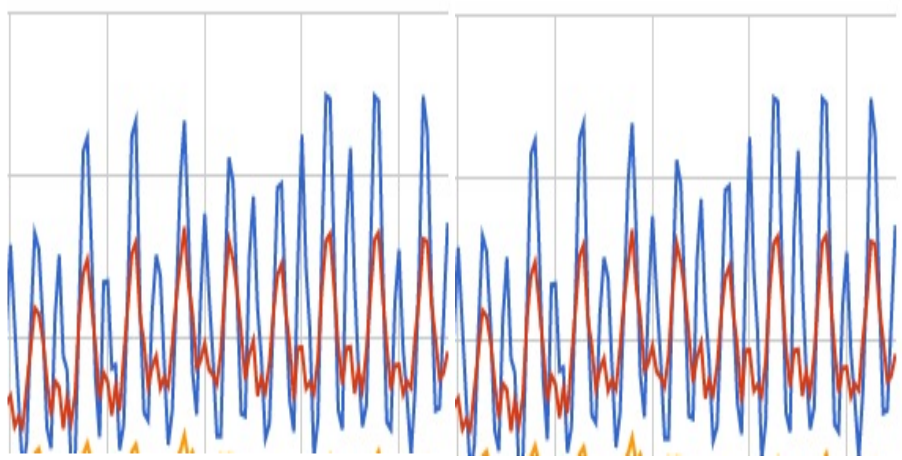


Temporal properties of variability: Relative Roughness

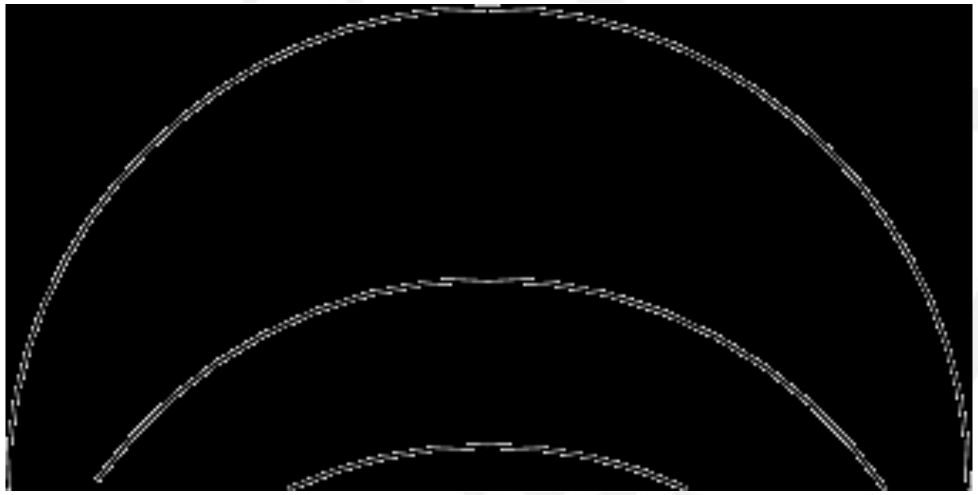
Relative roughness of a time series is:

$$RR = 2 * \left(1 - \frac{\text{local variance}}{\text{global variance}}\right)$$

Local variance:
Fast changes



Global variance:
Slow changes



Temporal properties of variability: Relative Roughness

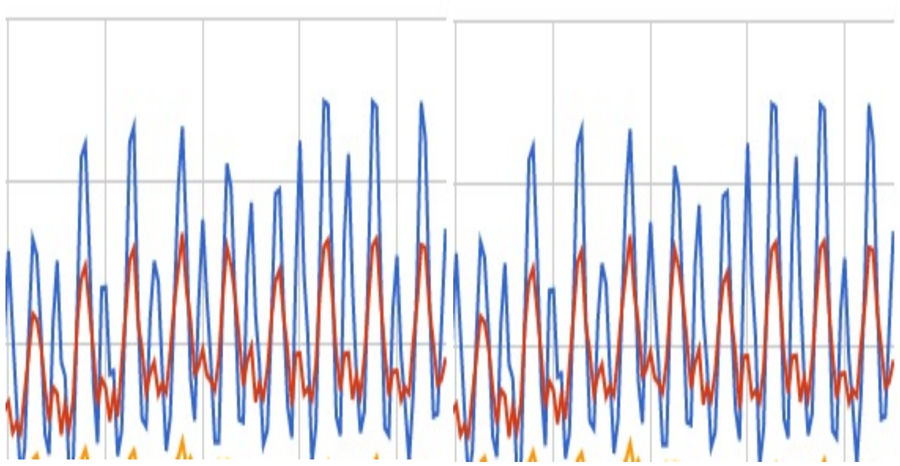
Relative roughness of a time series is:

$$RR = 2\left[1 - \frac{\gamma_1(x_i)}{\mathrm{Var}(x_i)}\right]$$

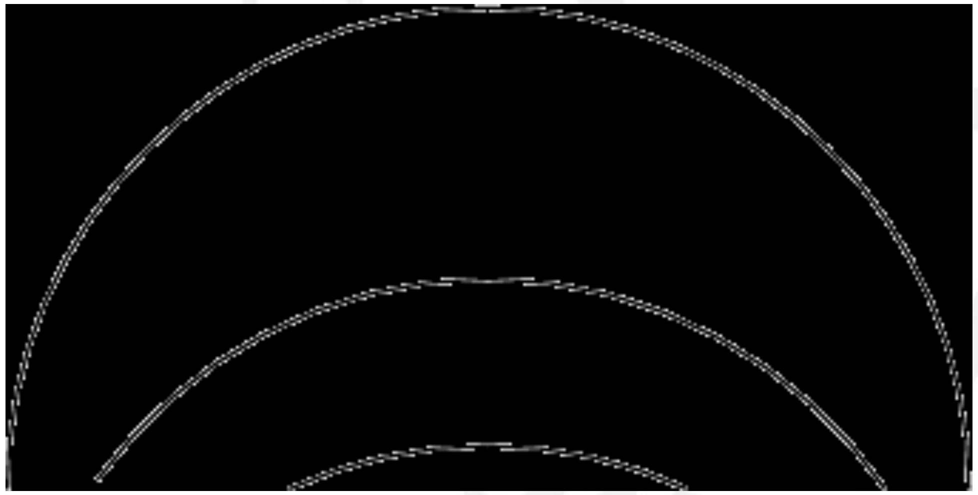
Lag 1 auto-(co)variance

Overall variance

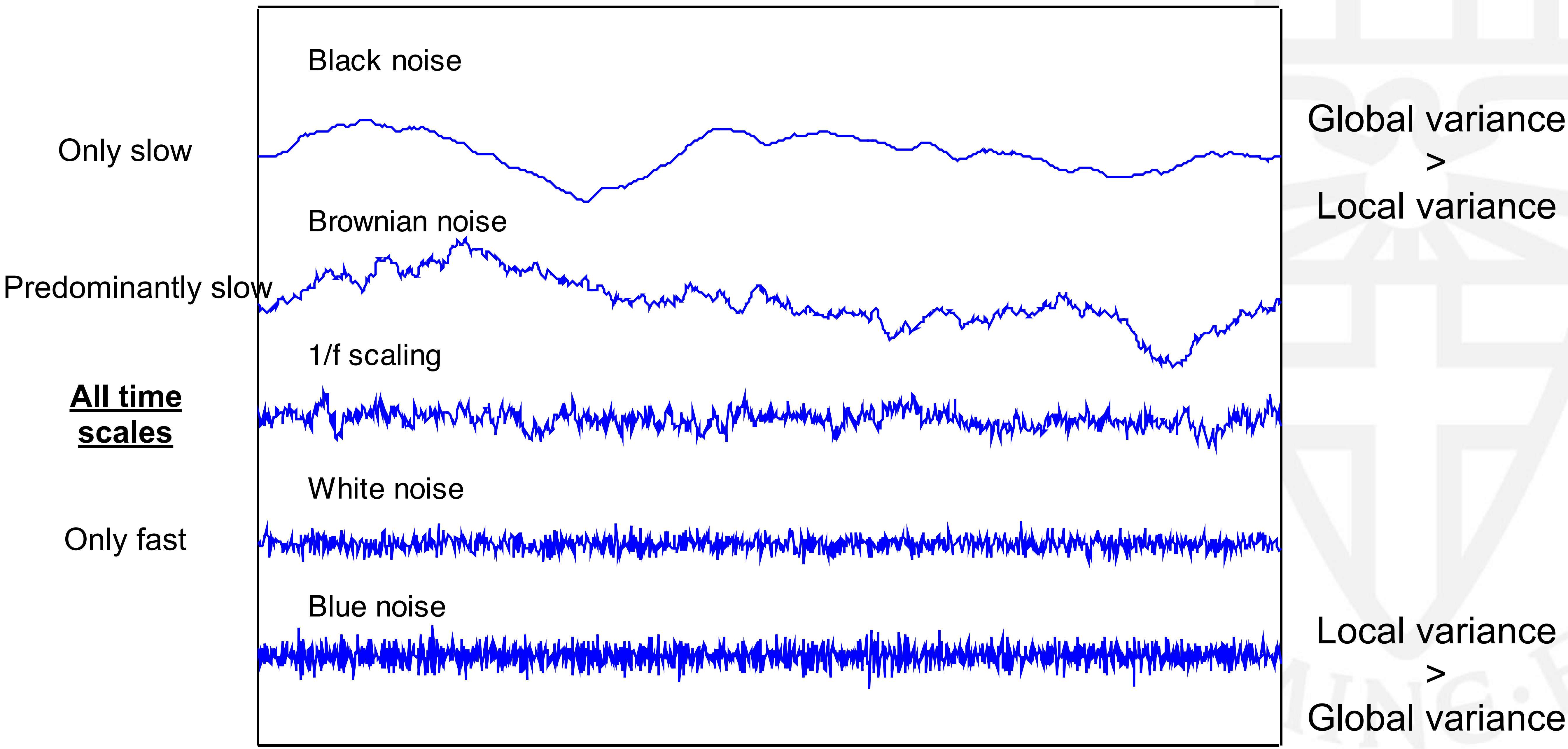
Local variance:
Fast changes

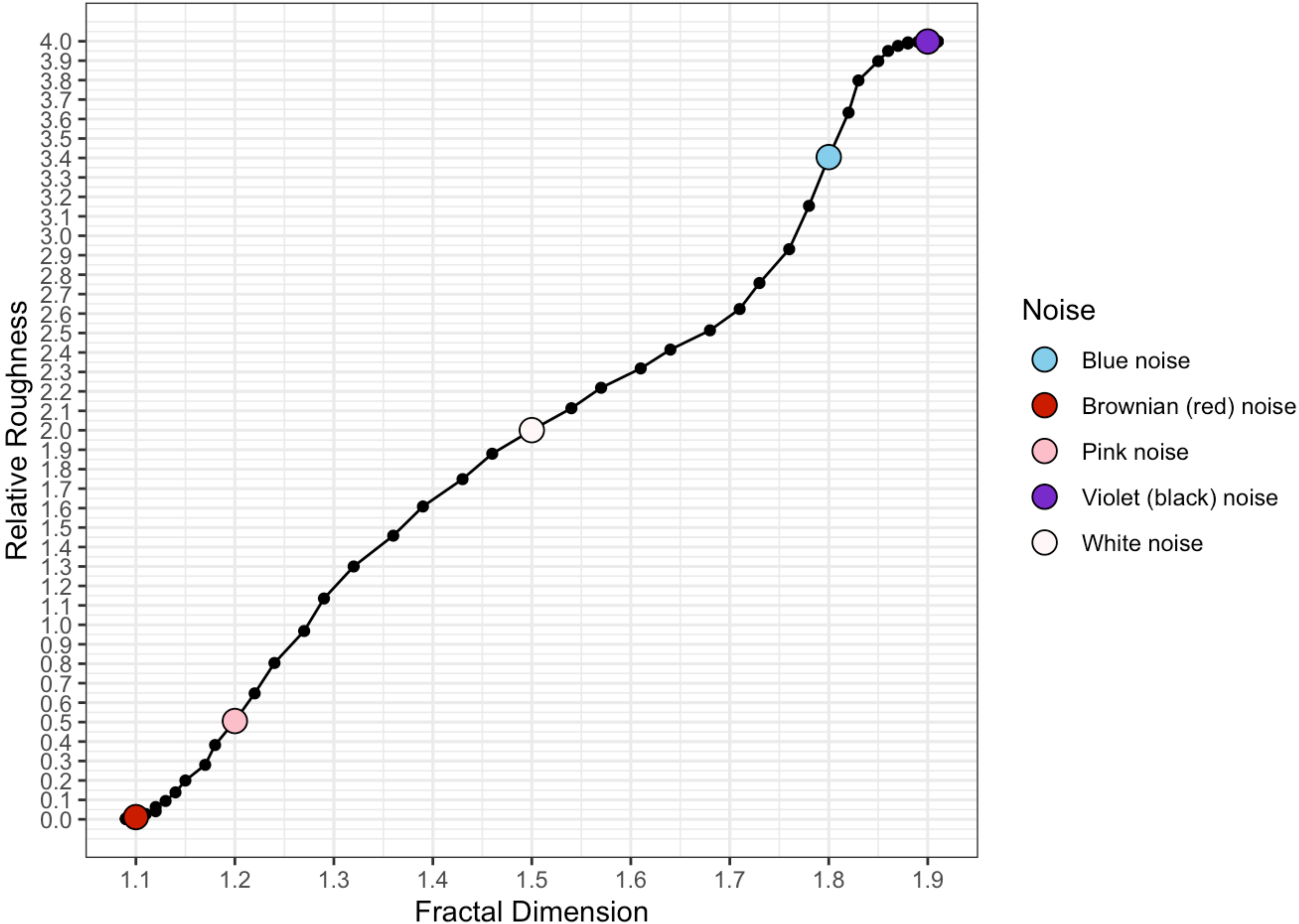


Global variance:
slow changes



Temporal properties of variability: Relative Roughness





Relative Roughness
4.0
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3.0
2.9
2.8
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2.0
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
Fractal Dimension
Noise
Blue noise
Brownian (red) noise
Pink noise
Violet (black) noise
White noise

Entropy

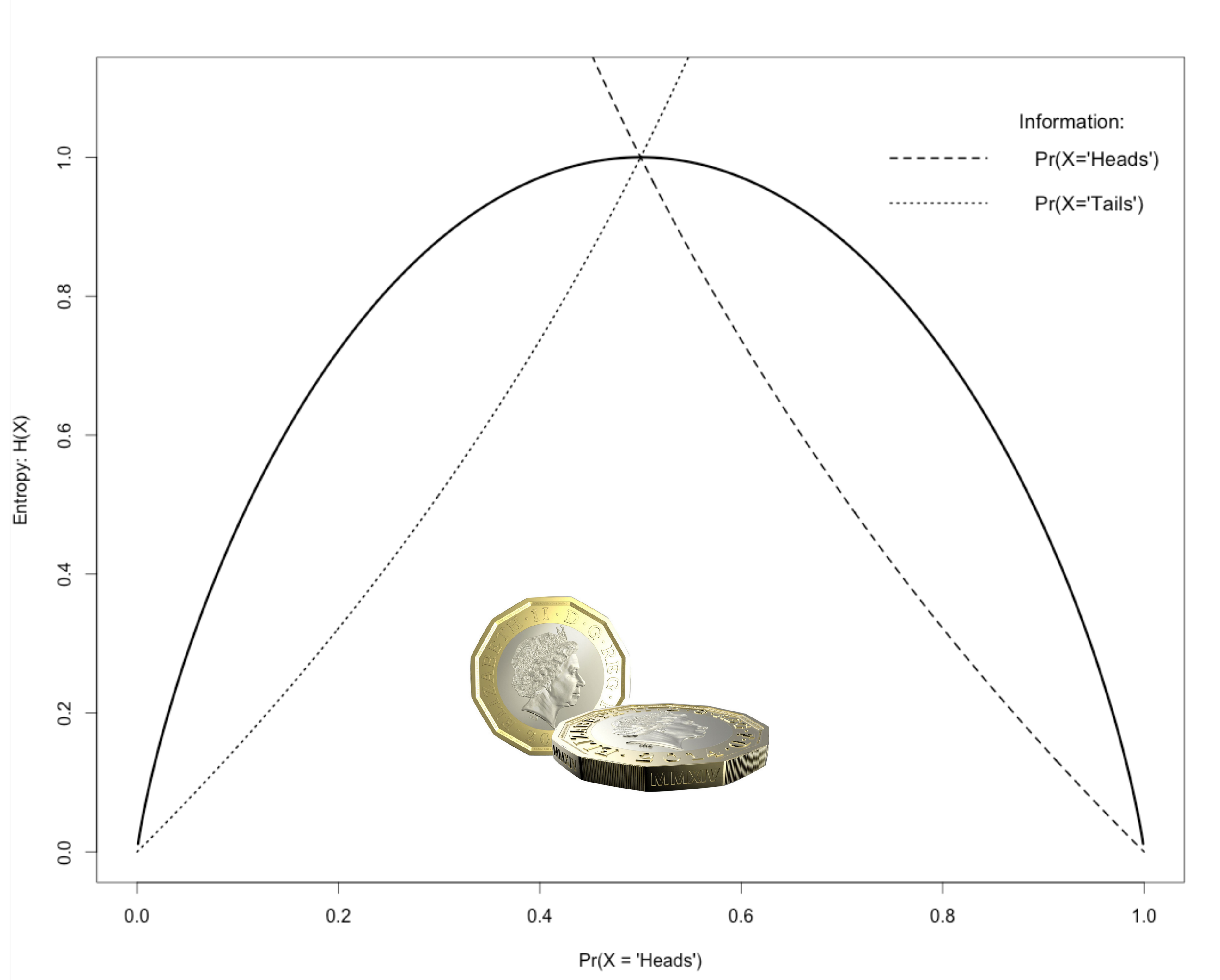
Entropy as a complexity measure

No obvious link with Roughness

- Different way to tap into dynamics

Entropy is a probabilistic measure of:

- uncertainty
- irregularity
- Predictability
- Information



Entropy

Information

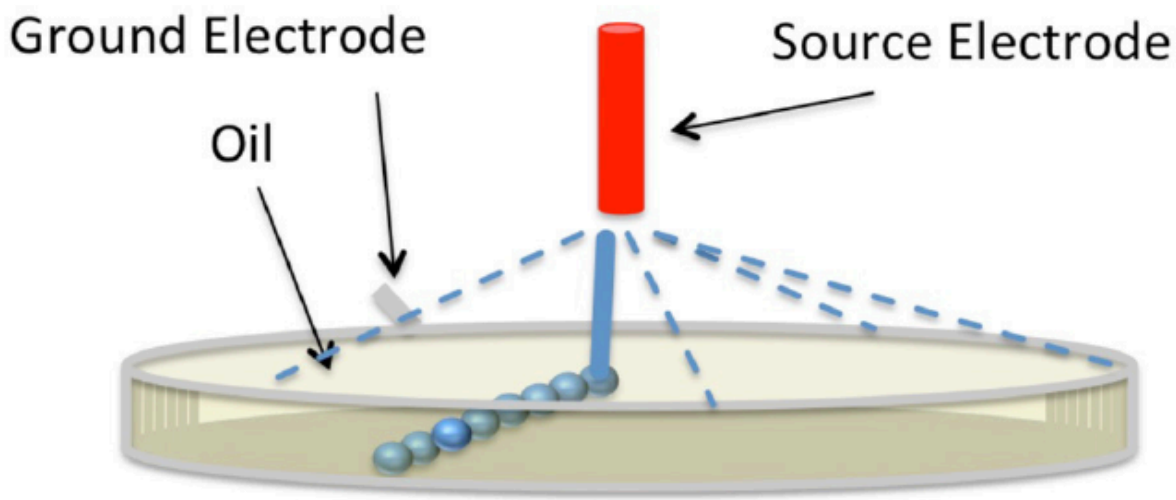
Uncertainty

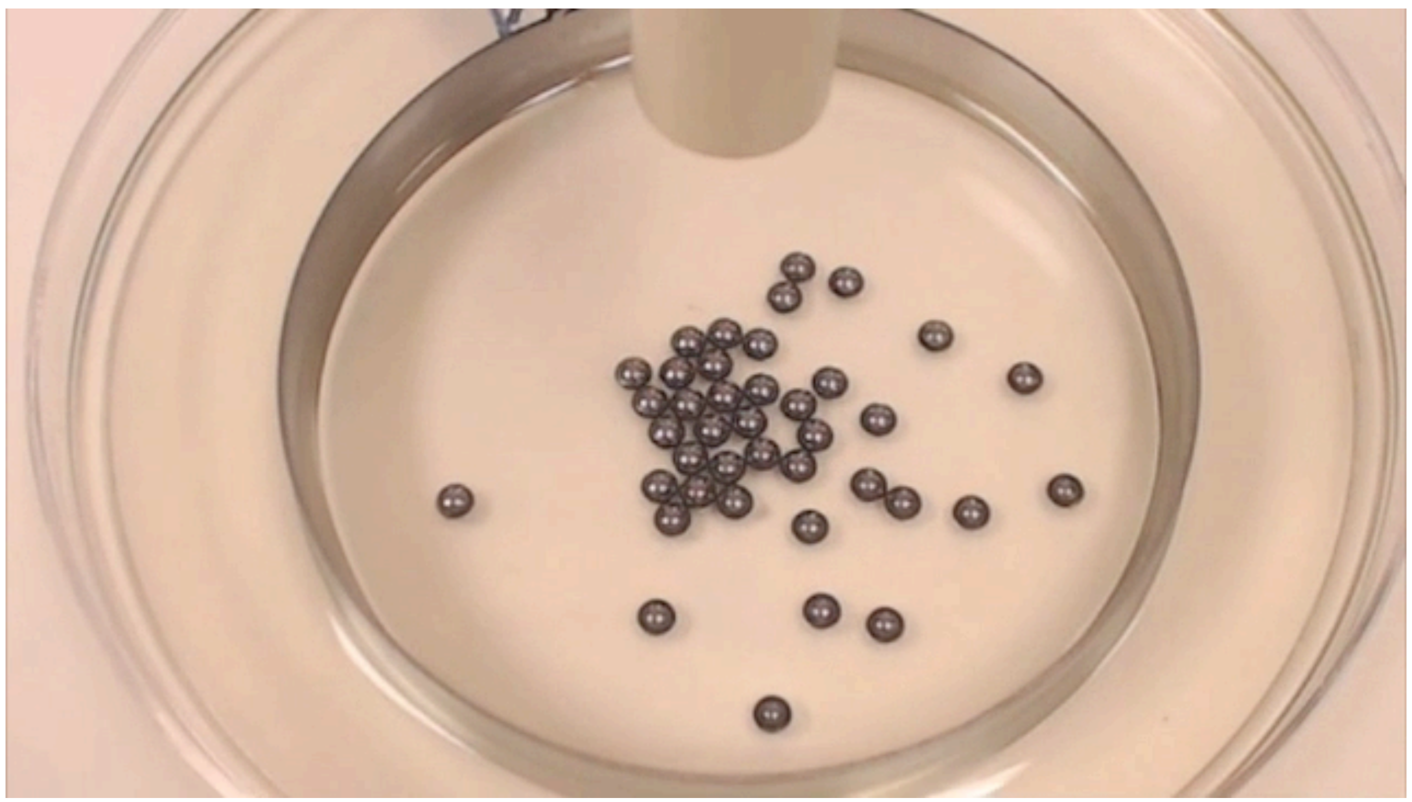
Redundancy

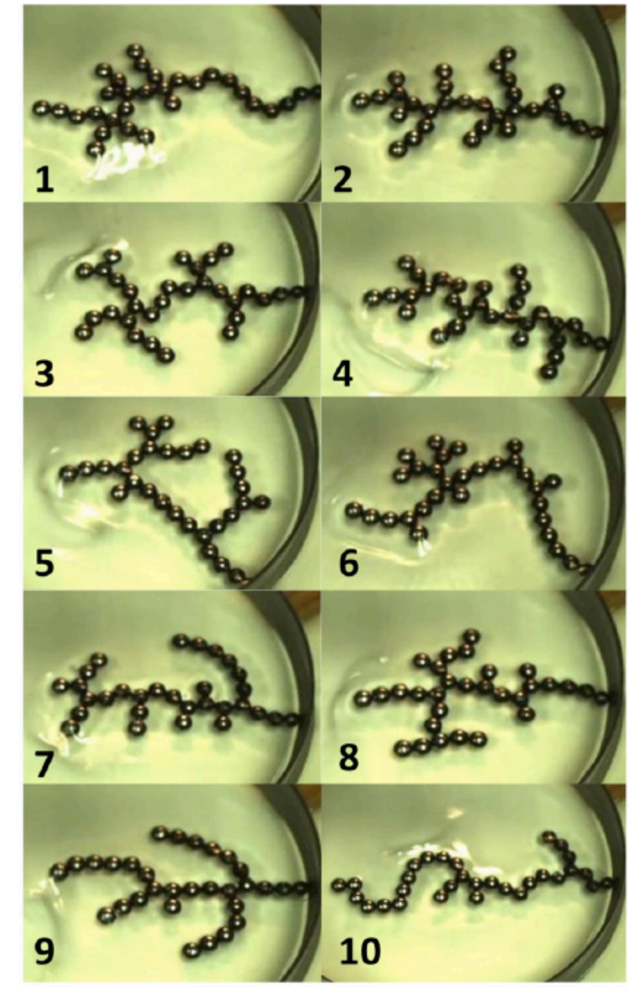
Probability

Symmetry breaking

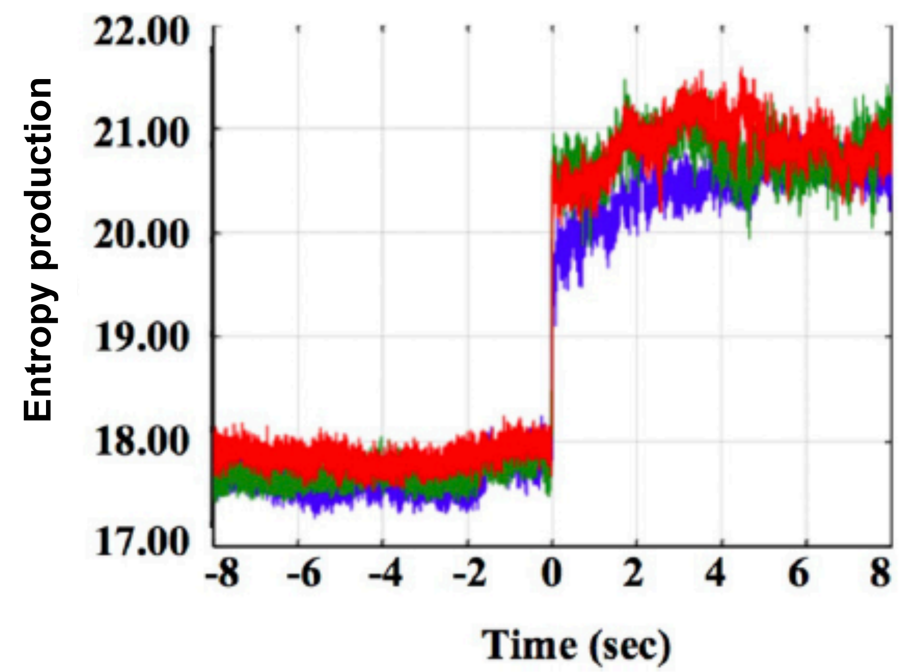
Complex behaviour from (physical) principles & lav







self-organisation:
Tree formation



Kondepudi D, Kay B, Dixon J. (2017). Dissipative structures, machines, and organisms: A perspective. *Chaos, 27(10),* 104607.

Disorder: Entropy State vs. Production

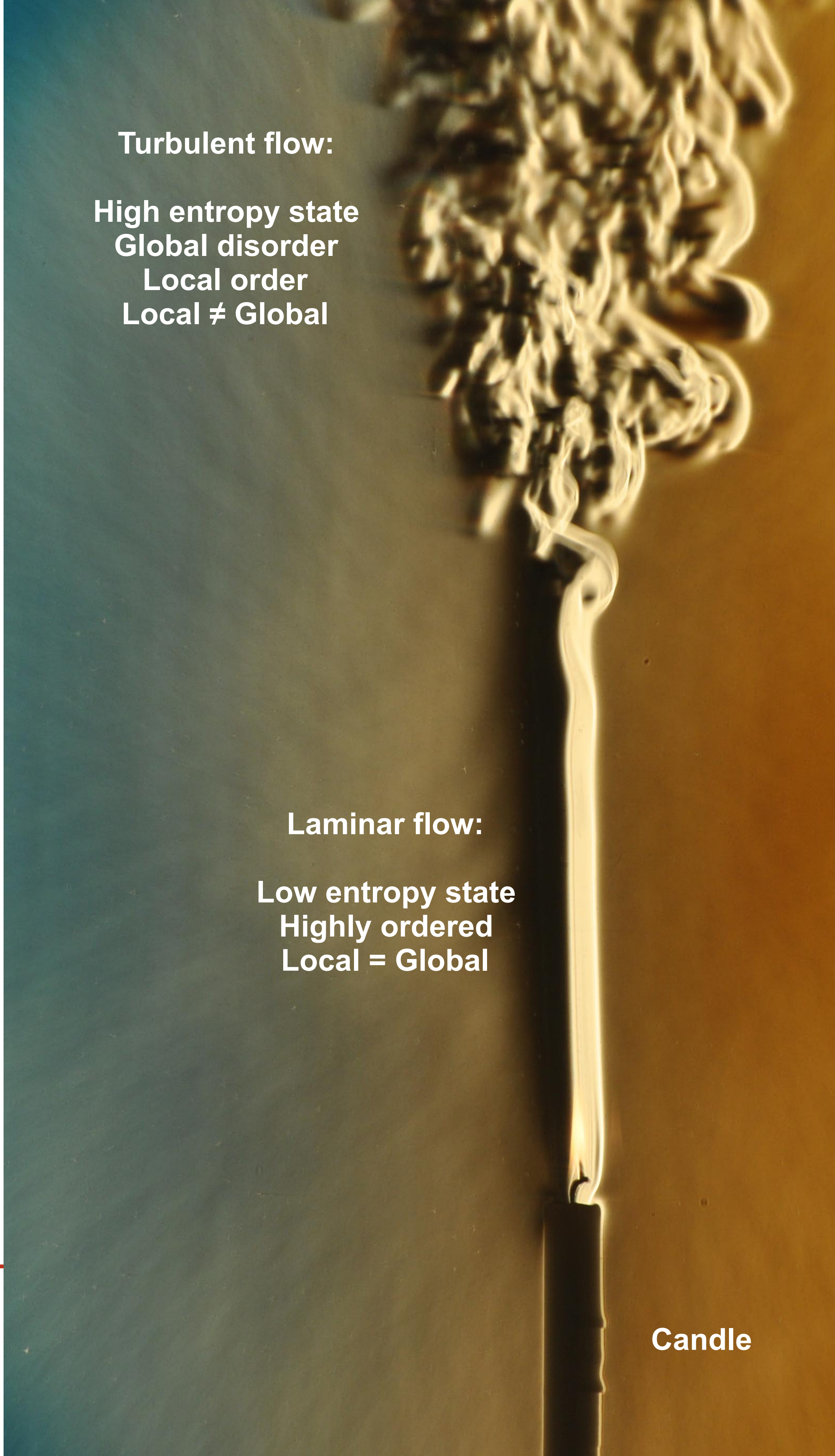
"**Modern thermodynamics is a theory of irreversible processes.** It characterizes irreversible processes into thermodynamic forces, F_k, that drive thermodynamic flows, J_k"

[…]

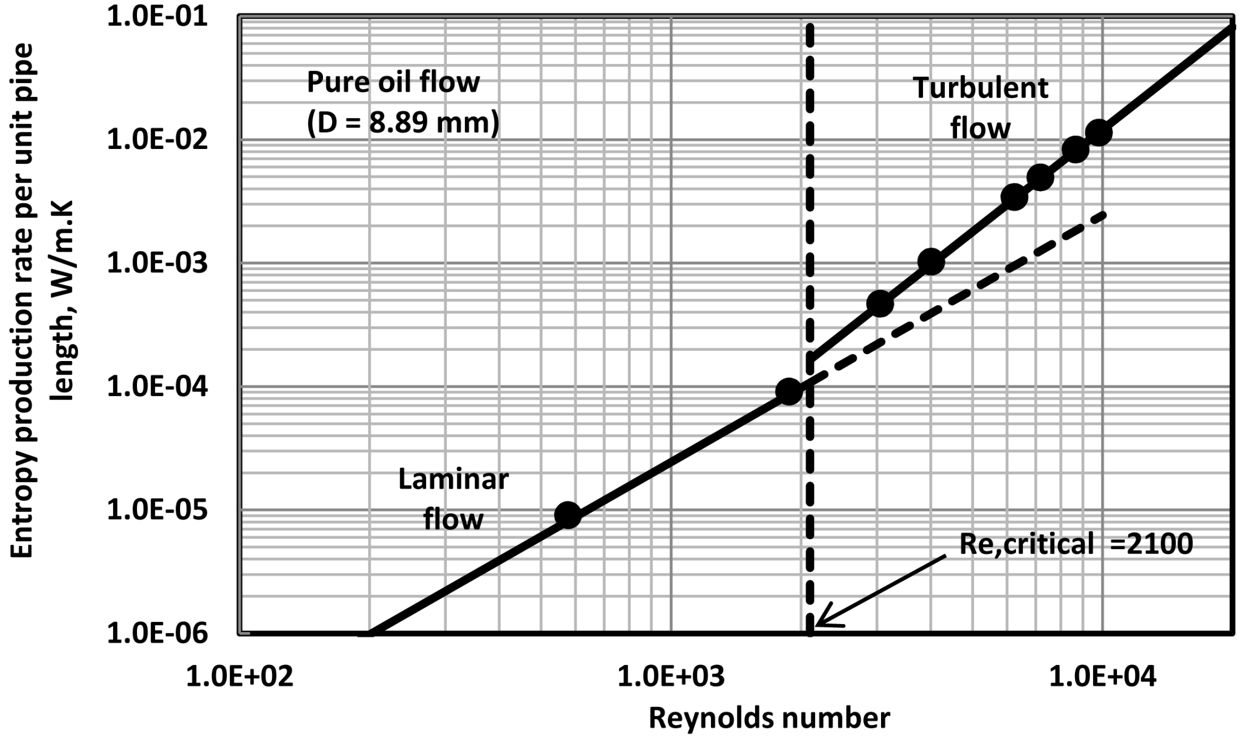
"All isolated systems, driven by entropy producing irreversible processes, ultimately reach a state of maximum entropy, the state of thermal equilibrium. At thermal equilibrium, all thermodynamic forces and thermodynamic flows that they drive vanish; there are no processes taking place in the system […] **they have produced all the entropy that can be produced in that system**" (Kondepudi, 2012)

Rate of Entropy Production (REP):

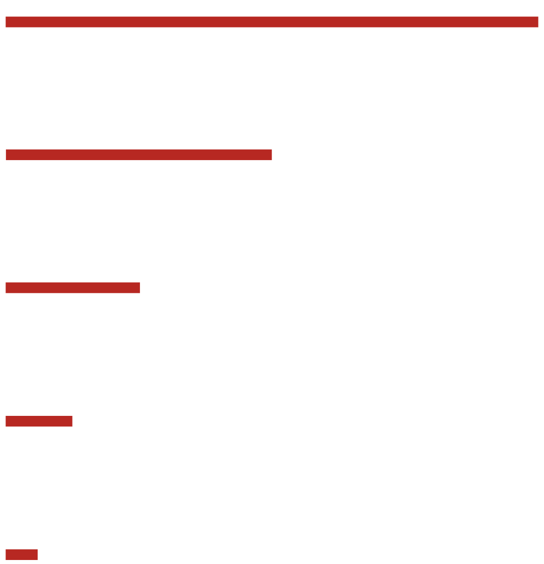
$$\sigma(x,t) = \sum_k F_k * J_k \geq 0$$



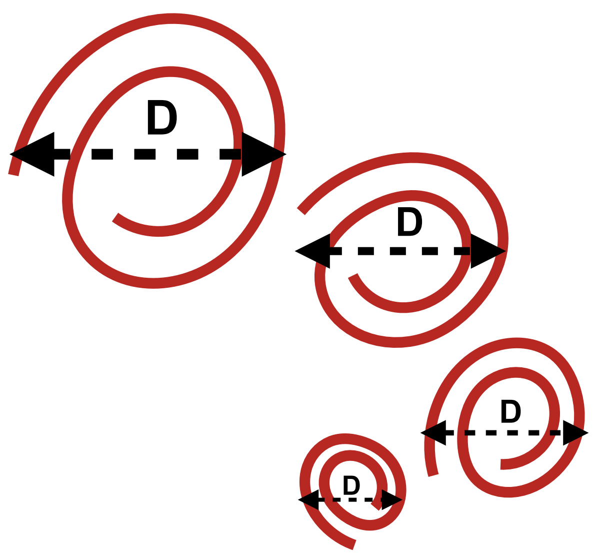
Disorder: Entropy State vs. Production

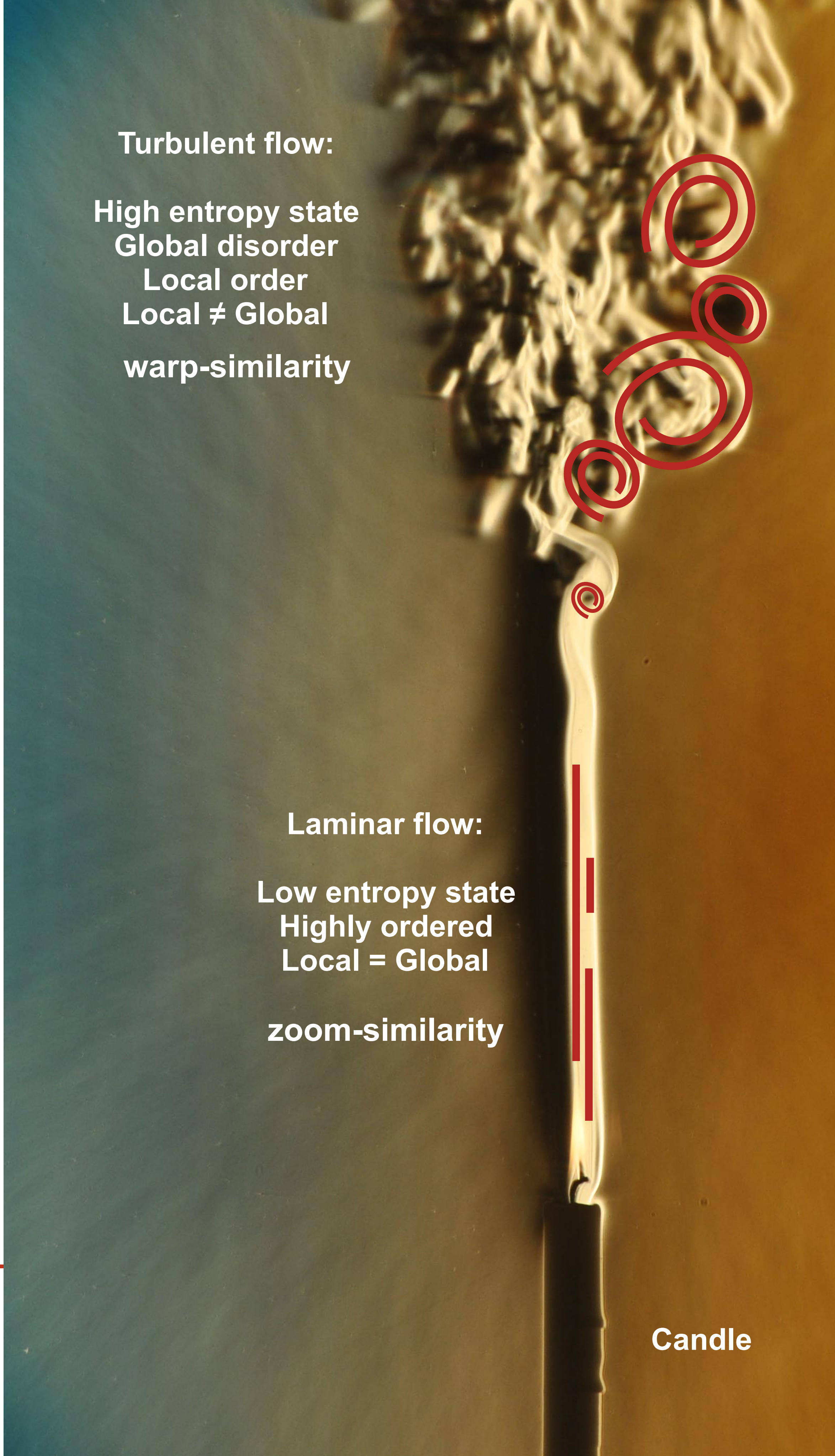


Linear/additive relation between scale and kinetic energy



Scaling relation between D and kinetic energy 5/3
Kolmogorov's Constant

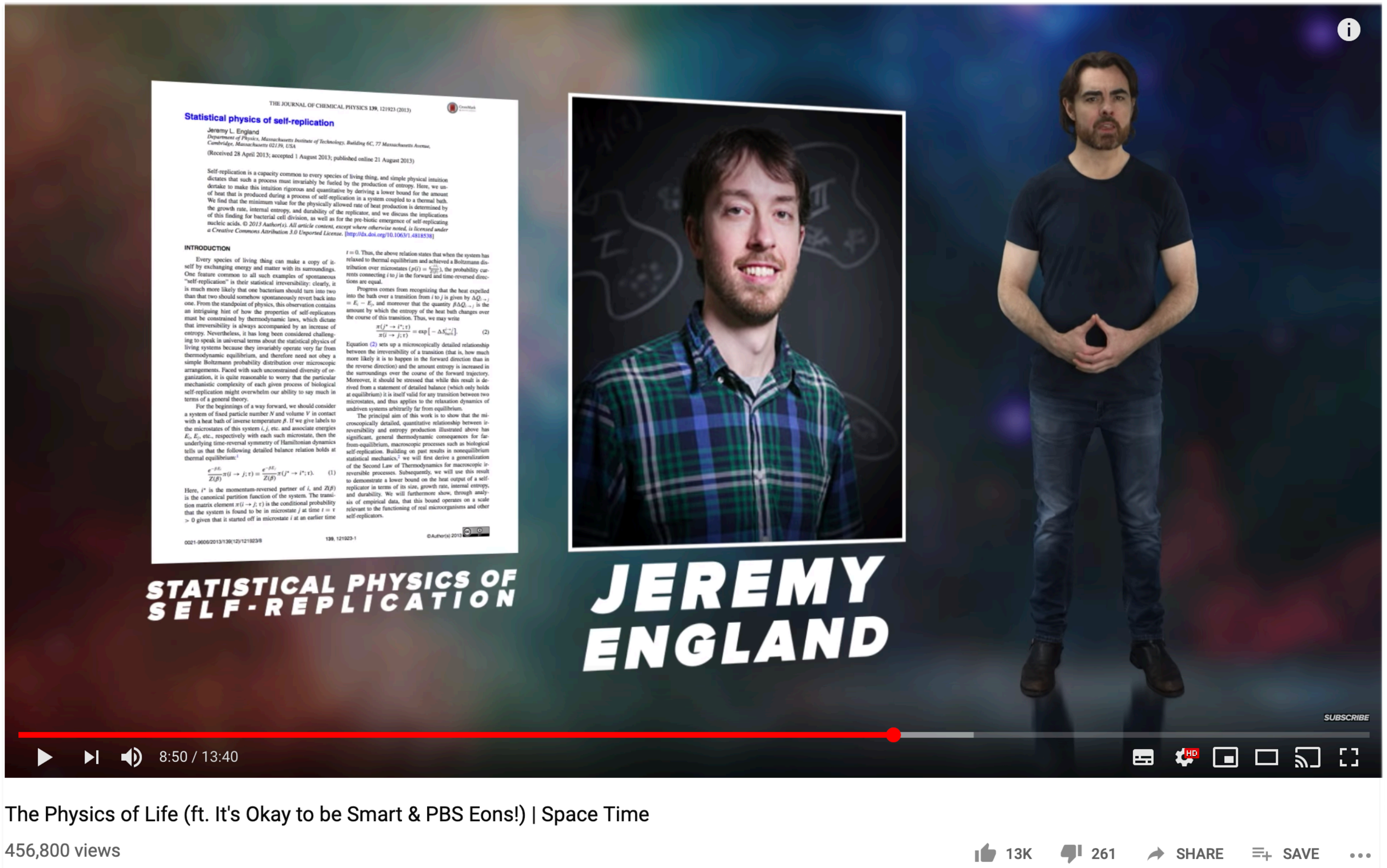




https://youtu.be/GcfLZSL7YGw

Natural Computation in Physics >> Dissipative Systems

"*… regard the physical world as made of information, with energy and matter as incidentals*" -Bekenstein (2003, p.59)

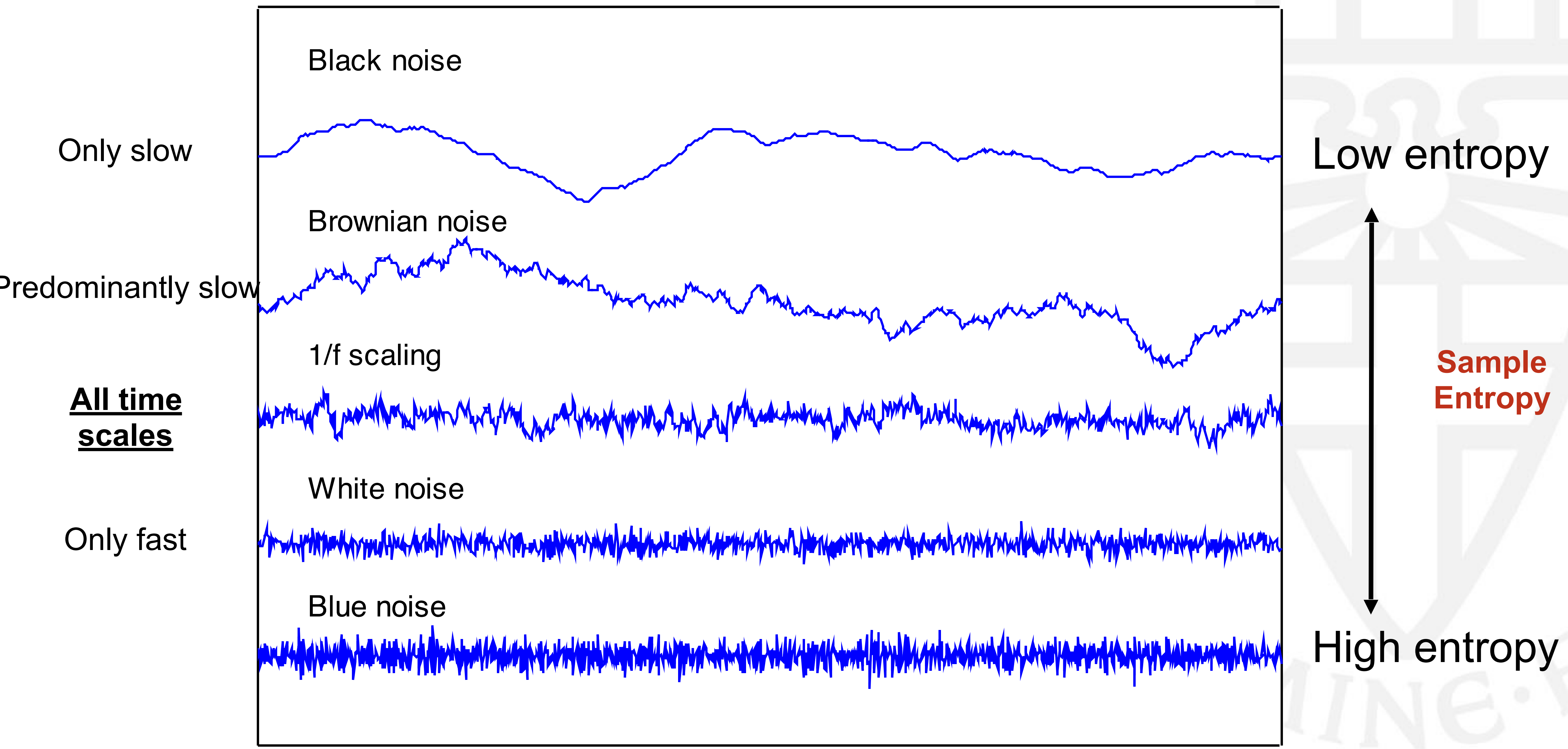


Physics of Life - How does complexity emerge under 2nd Law of Thermodynamics?

- Open systems >> Continuous flow of energy >> Energy gradients >> Pattern formation
- High energy / ordered states need to be dissipated as heat / disorder (2nd Law)
- However, stable patterns emerge that eventually self-replicate >> Natural selection mechanism
- Self-replicating systems turn out to be efficient order / energy dissipators! (e.g. exponential growth)

England, J. L. (2013). Statistical physics of self-replication. *The Journal of chemical physics*, *139*(12), 121923. https://doi.org/10.1063/1.4818538

Temporal properties of variability: Sample entropy



Entropy in time series data

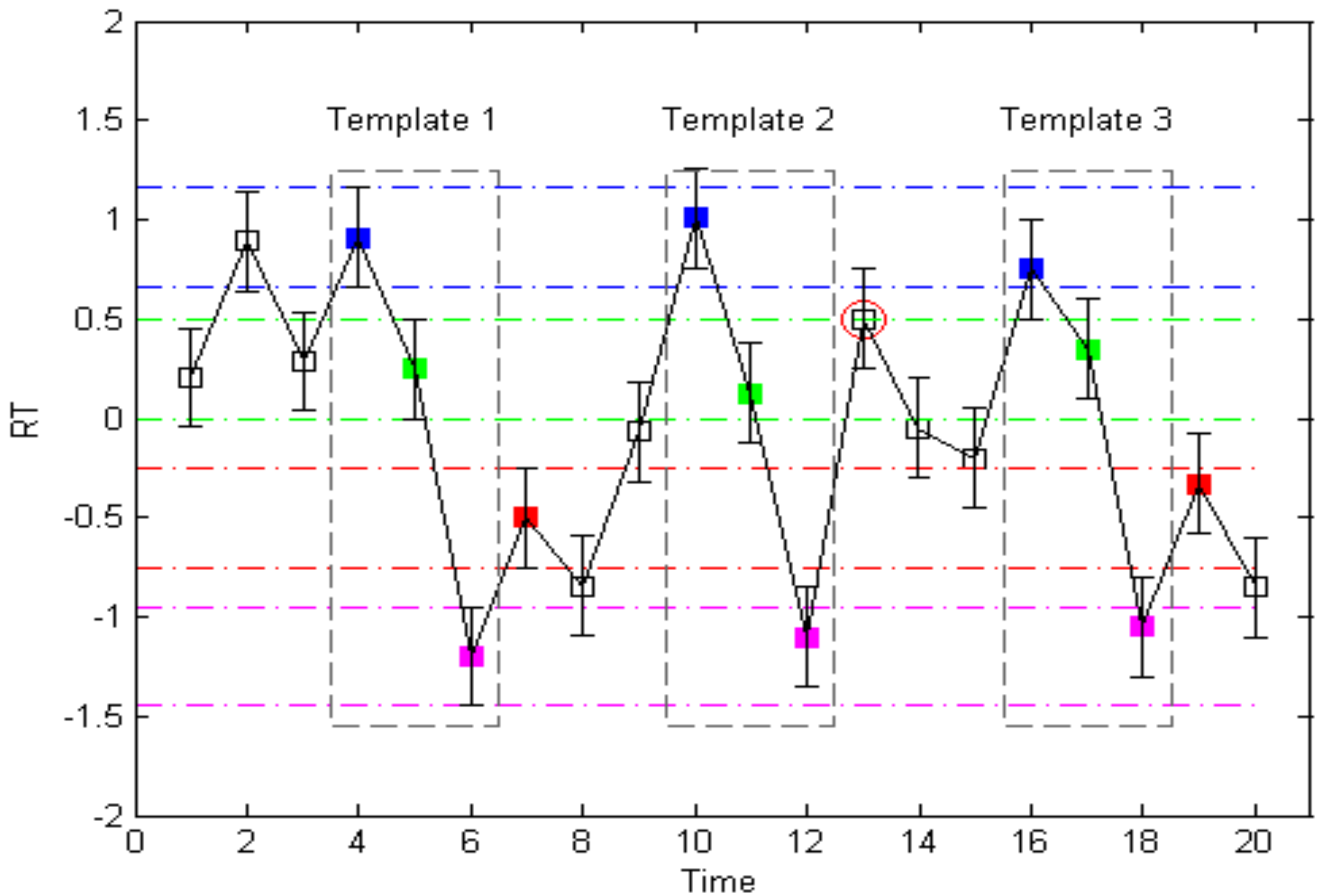
Sample entropy

- P = $A(k)/B(k)$
  - A: # of data segment of length m+1 are within distance < r
  - *B: #* of data segment of length m are within distance < r

- The negative natural logarithm of the conditional probability that a dataset of length N, having repeated itself within a tolerance r for m points, will also repeat itself for $m + 1$ points.

- SampEn $(m, r, N) = -\ln P$

- SampEn: the negative natural log (-ln) of the conditional probability that the pattern of *m*+1 points (■-■-■-■) will match if a pattern of *m* points (■-■-■) did match



Sample entropy

Determine *m*

- the length of compared runs of data
- E.g., 3 data points

Determine *r*

- Tolerance range
- E.g., 1 standard deviation

Sample entropy: interpretation

A small value (e.g., 0.05)

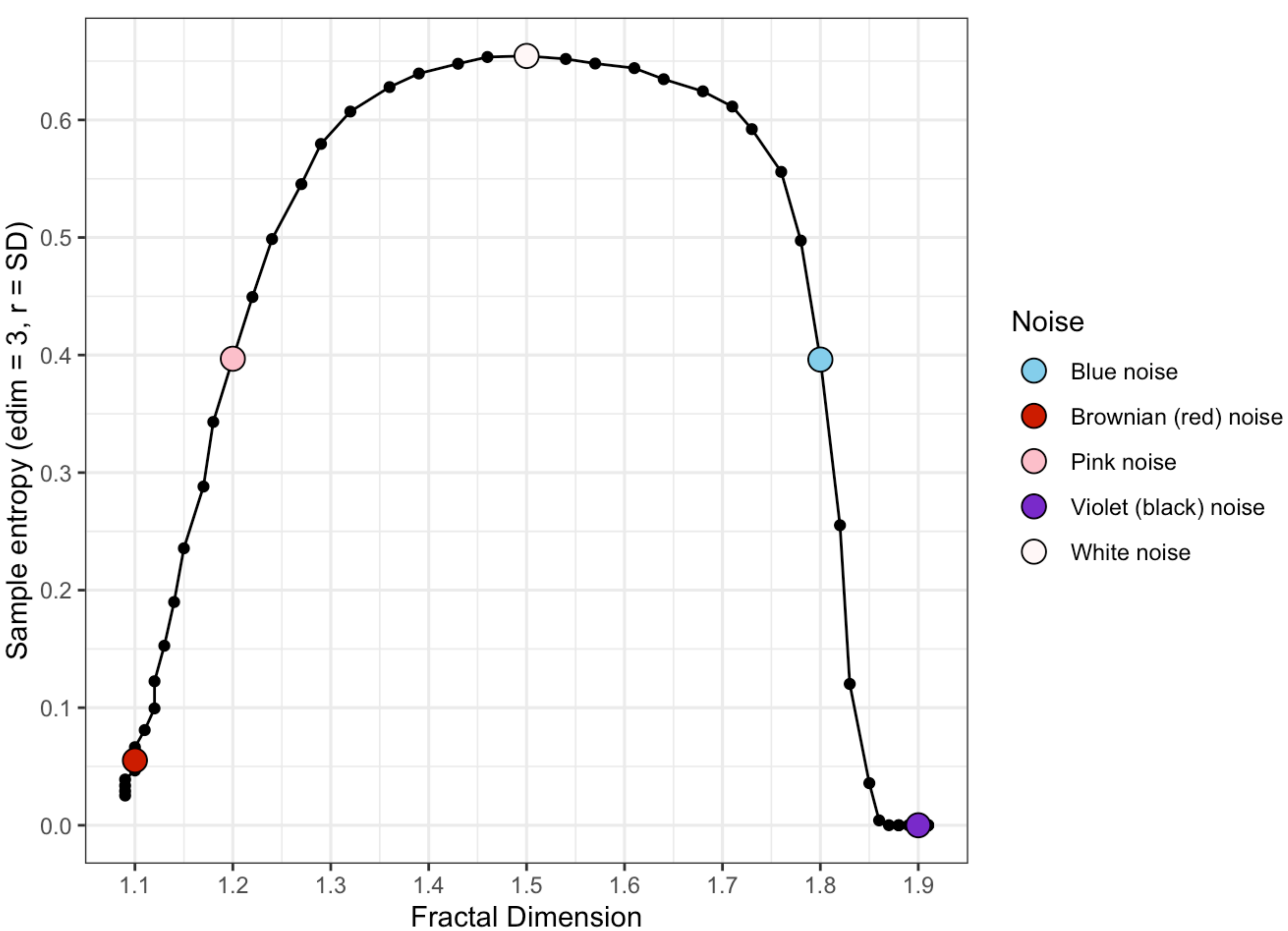
- sequence is regular and predictable
- a high probability of repeated template sequences in the data

A large value (e.g., 1.5)

- sequence is irregular and unpredictable
- a low probability of repeated template sequences in the data

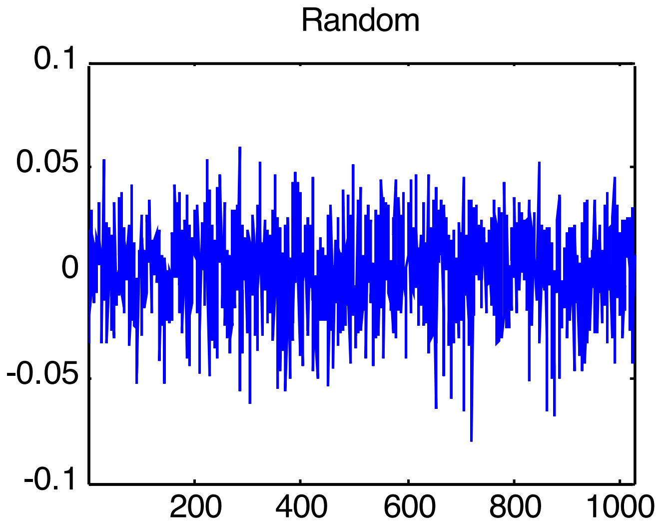
NOTE: absolute values will change in function of your parameter choices for *m* and *r*

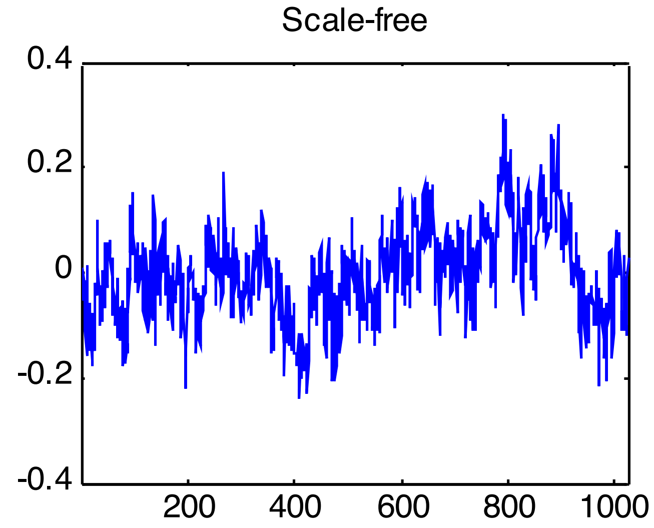
- the number of matches can be increased by choosing small *m* (short templates) and large *r*(wide tolerance).

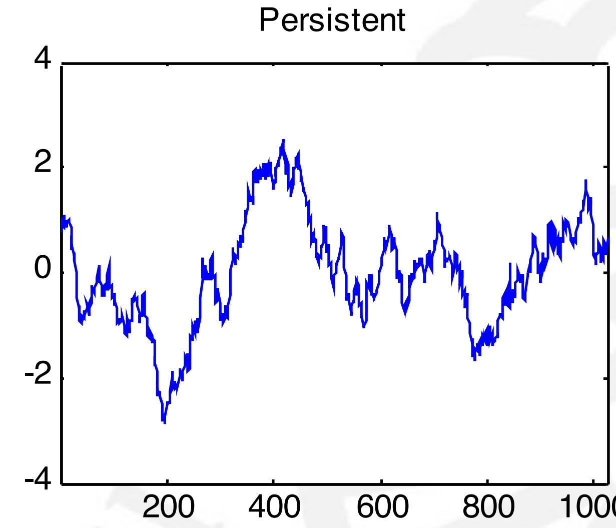


Sample entropy (edim = 3, r = SD)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
Fractal Dimension
Noise
Blue noise
Brownian (red) noise
Pink noise
Violet (black) noise
White noise

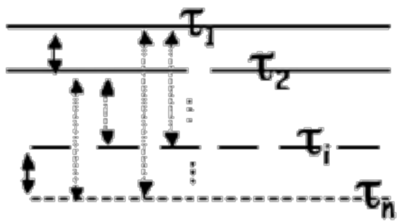
Time series analysis: sum up





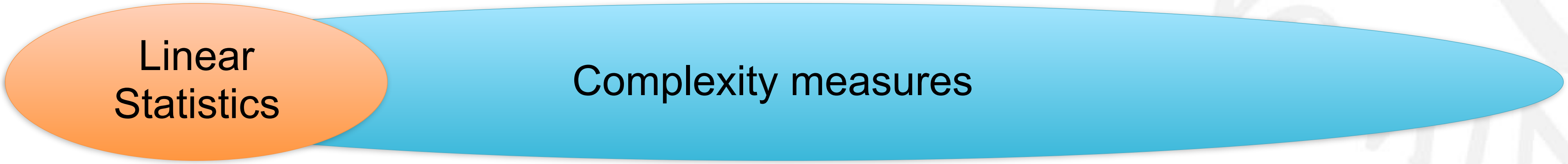


- Flexible
- Disorganized
- No slow time scales
- Unconstrained
- Many degrees-of-freedom



Dynamics at all
time scales contribute
to the process

- Rigid
- Order
- Predominantly slow time scales
- Constrained
- Few degrees-of-freedom



Scaling Phenomena

EXTRA

Scaling phenomena

How long is the coast of Great-Britain?

Scaling phenomena

scale

11.5 x 200 = 2300 km

scale

28 x 100 = 2800 km

scale

70 x 50 = 3500 km

Length <u>systematically</u> depends on the size of the measurement stick you use!

https://en.wikipedia.org/wiki/Fractal_dimension

Scaling phenomena

"scaling of bulk with size"

(Theiler, 1990)

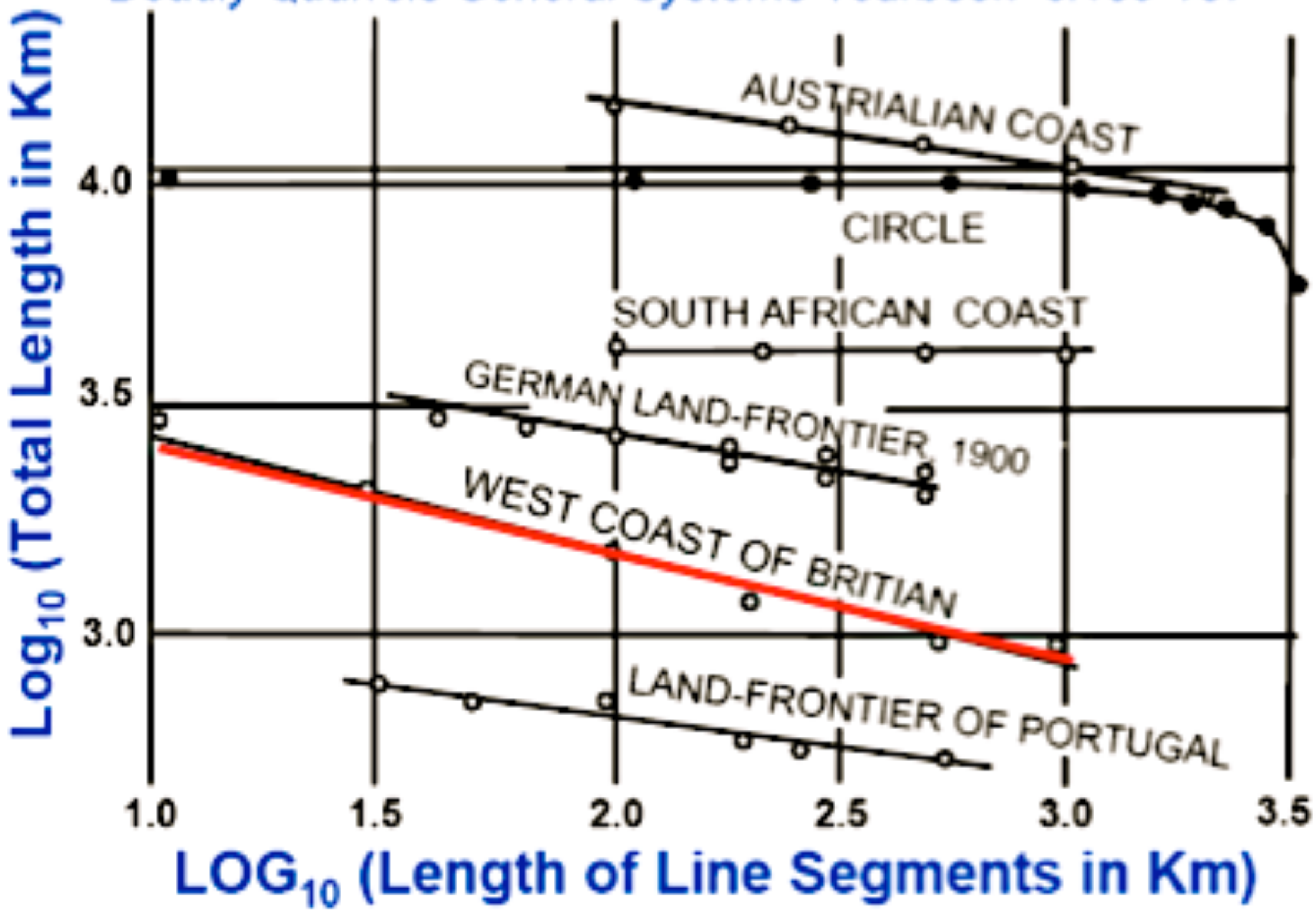
The formal answer to the question is:

"There is no characteristic scale at which the length of the coast of GB can be expressed"

Mandelbrot, B. B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science, 156*(3775), 636–8.

How Long is the Coastline of Britain?

Richardson 1961 The problem of contiguity: An Appendix to Statistics of Deadly Quarrels General Systems Yearbook 6:139-187



Scale invariance…

no meaningful central moments can be defined

Mean and SD characterise the data only relative to the scale of observation (e.g. sample size)

A *power law* scaling relation (**LOG scale**):
There is no characteristic length, just an indication of complexity

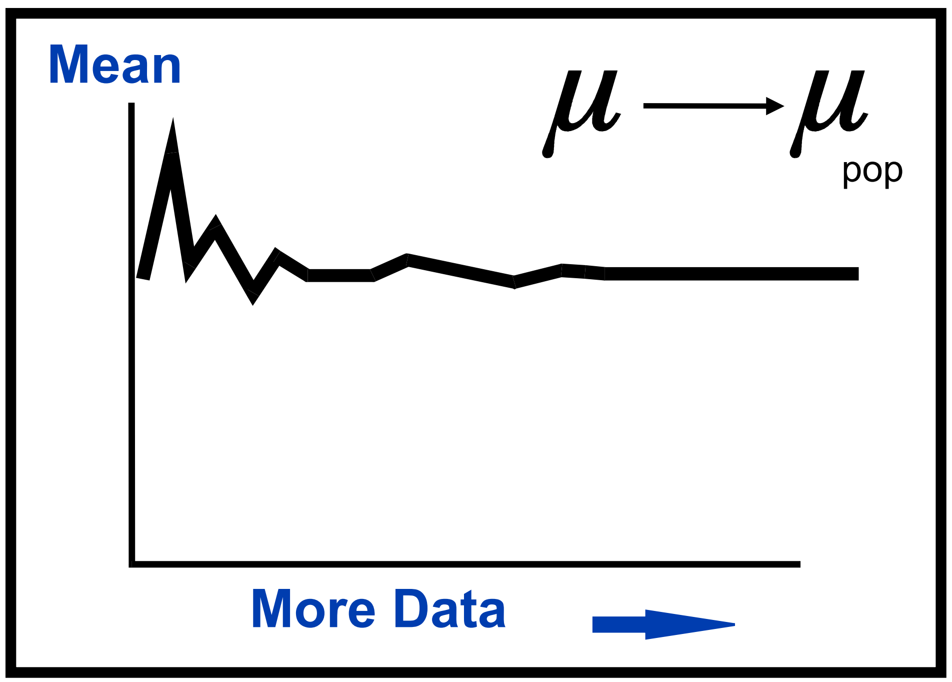
Scaling phenomena

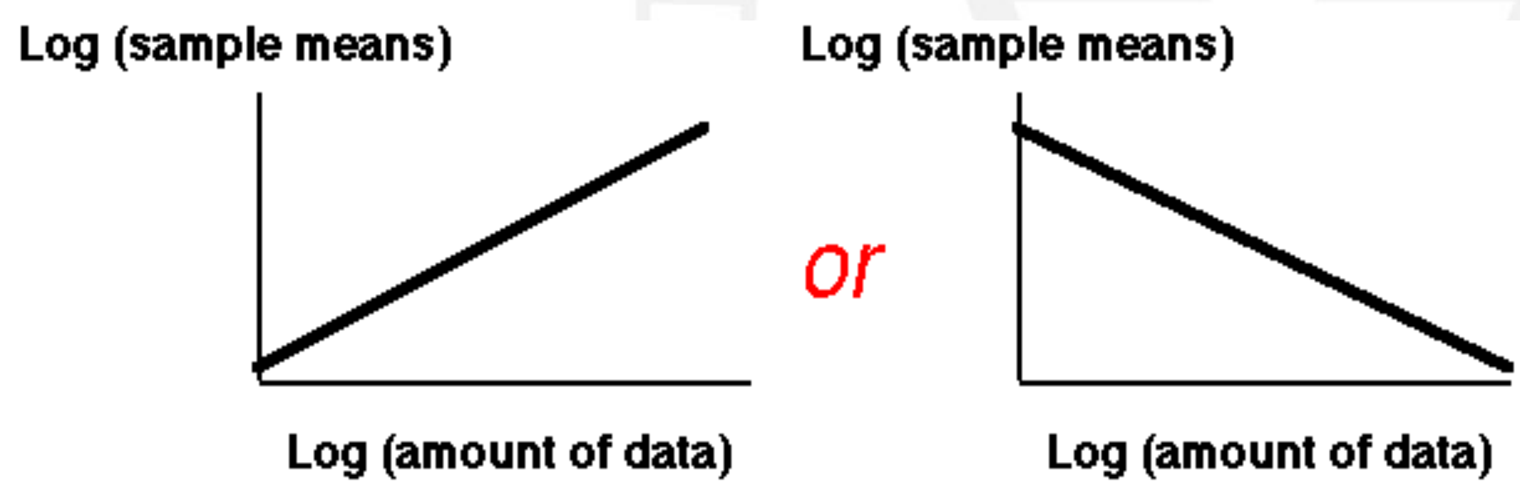
Independent observations of random variables

μ±σ are sufficient to characterise absence of dependencies in the data:
e.g. Expected value of **μ** for **N** = 100, given **σ**

Interdependent observations across different scales

μ±σ are insufficient to characterise dependencies in the data:
e.g. Sample estimates of **μ** change with **N**





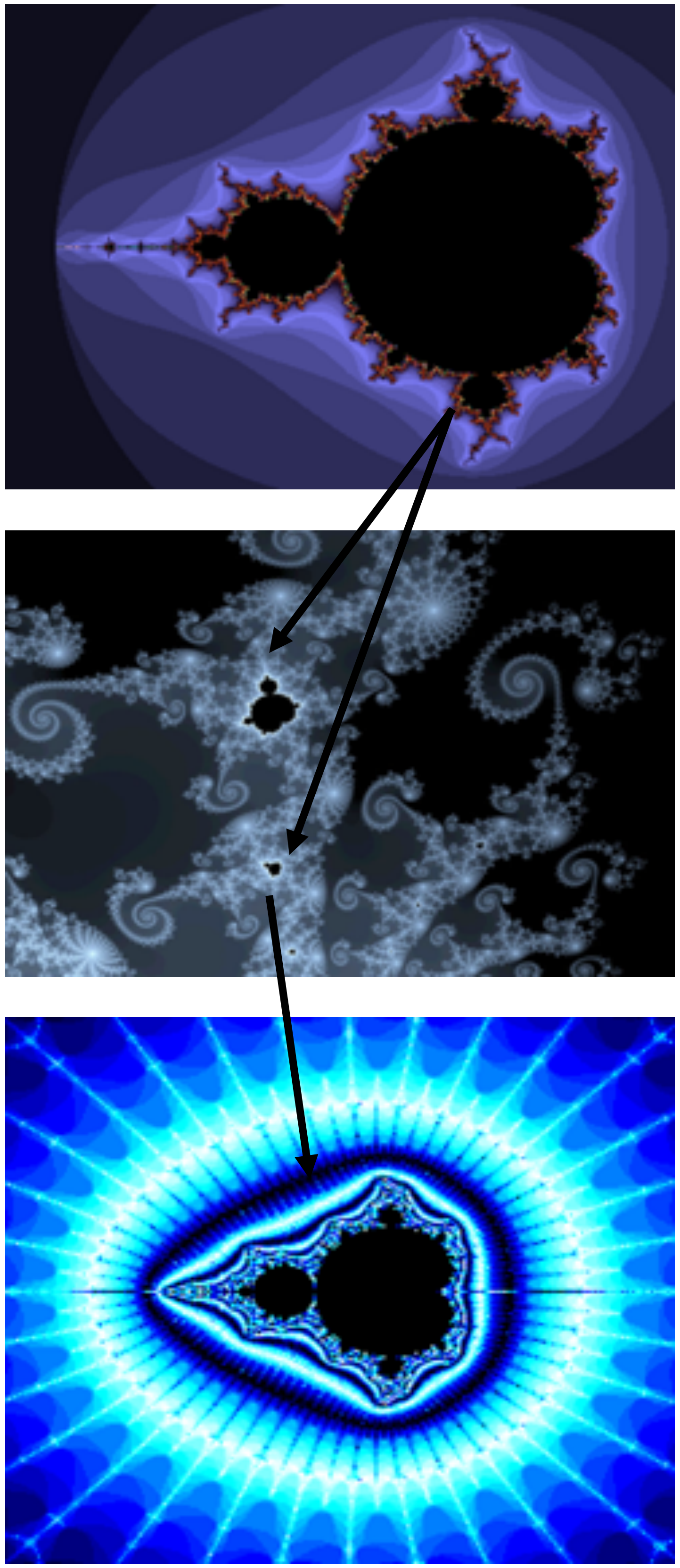
Help! How can I do science without **μ** or **σ**?

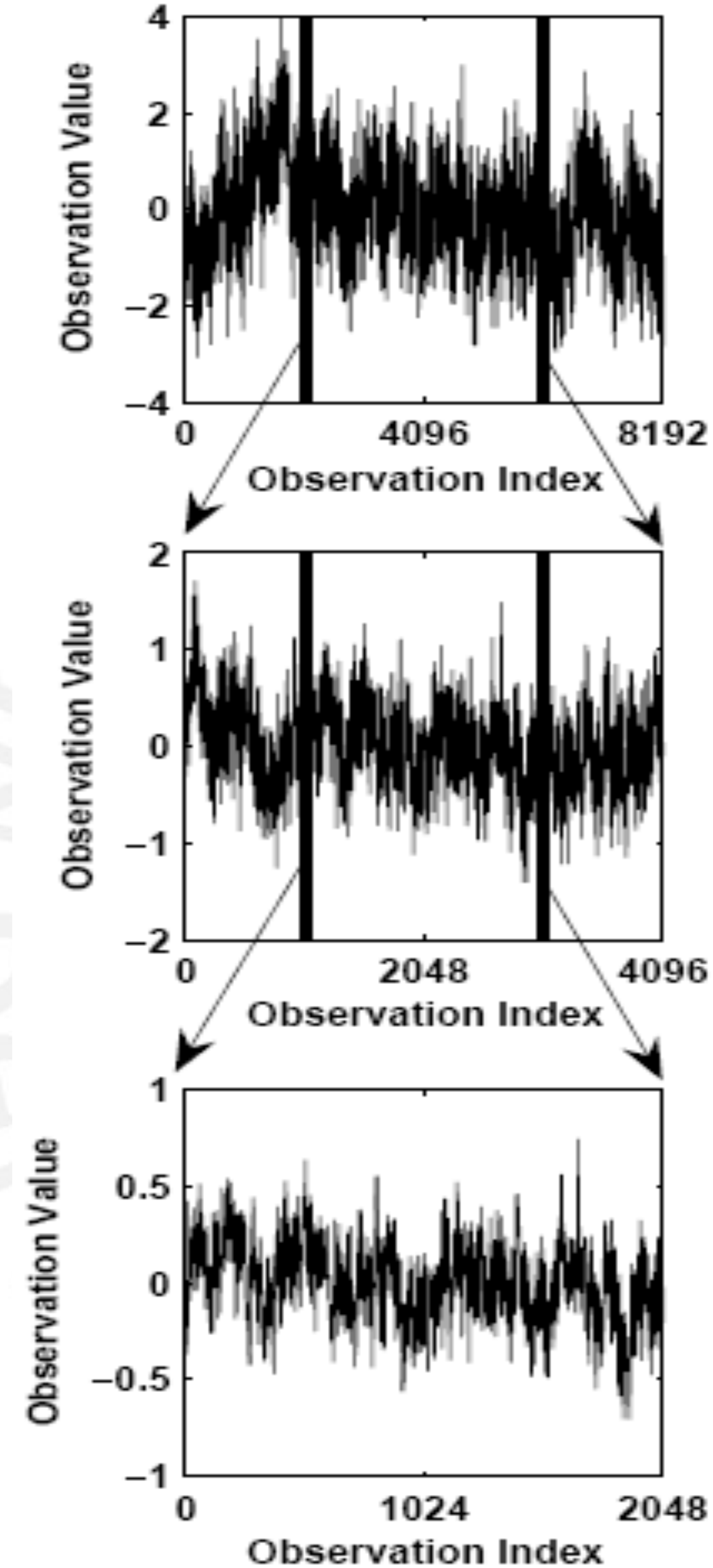
What is scaling?
Self-similarity and Self-affinity

Object looks roughly the same on all scales = (Statistical) **self-similarity** (“zoom similarity”)

(Statistical) self-similarity is observed after affine transformation = **self-affinity** (“warp similarity”)

Degree of invariance across scales = Dependencies/regularities/correlations across scales





aka: “Fractal scaling”

Scaling phenomena

Scaling relations can emerge with all kinds of observables
They inform about properties of the process / system under scrutiny

Earthquakes (Richter-Law)

frequency of occurrence ~ magnitude

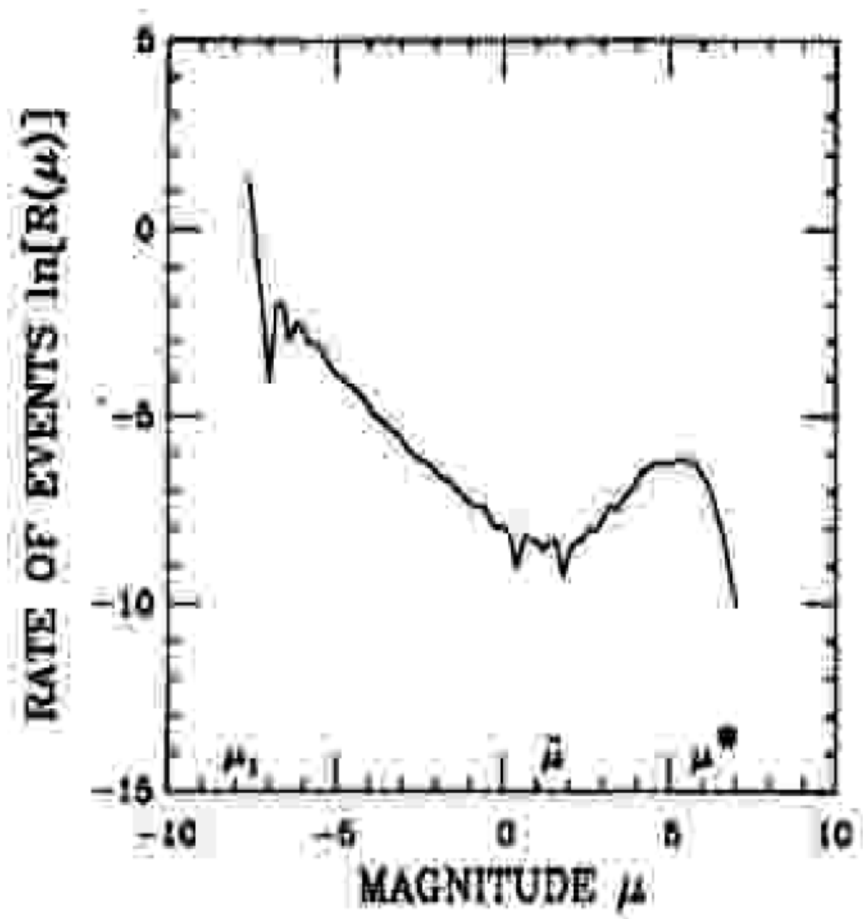
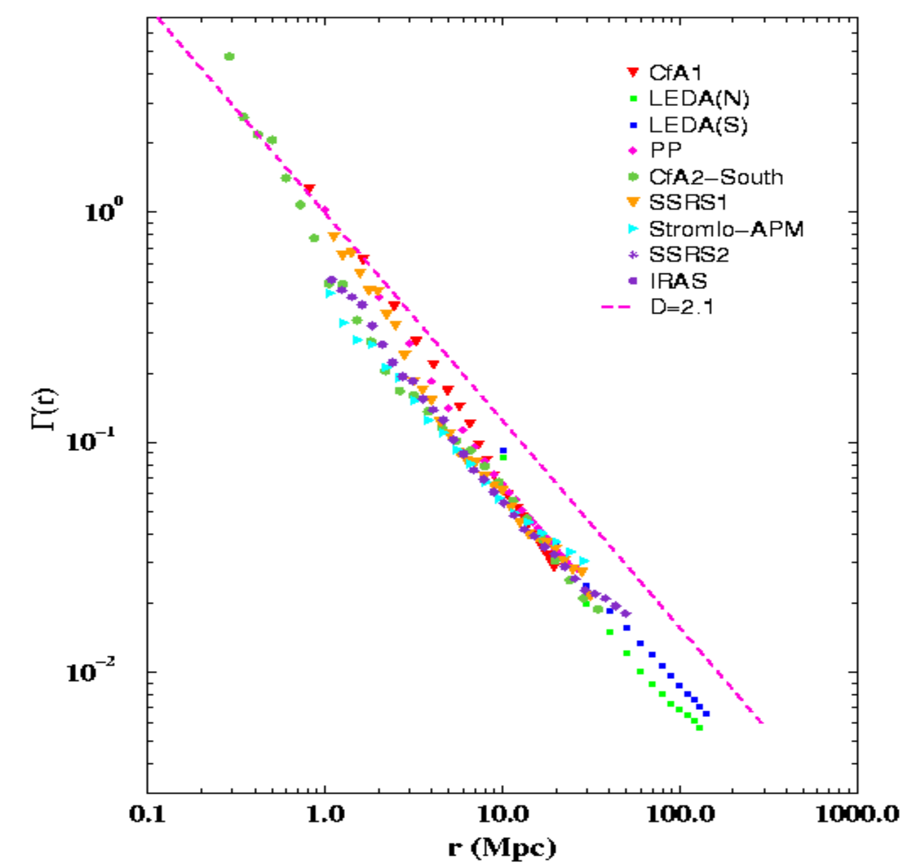


Figure 13: frequency distribution of the slip events (earthquakes) of magnitude μ taken from [53]. Notice the large bump that corresponds to an excess of events of high magnitude.

Distribution of mass in the Universe

resolution ~ density

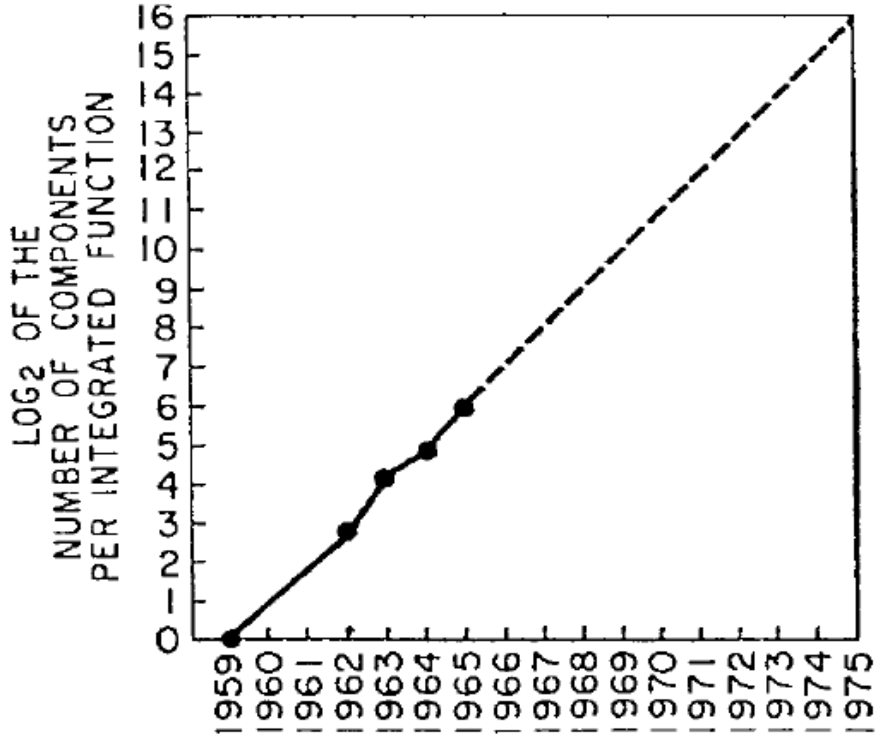


Behavioural Science Institute
Radboud University Nijmegen

Scaling & Growth

Moore's Law:

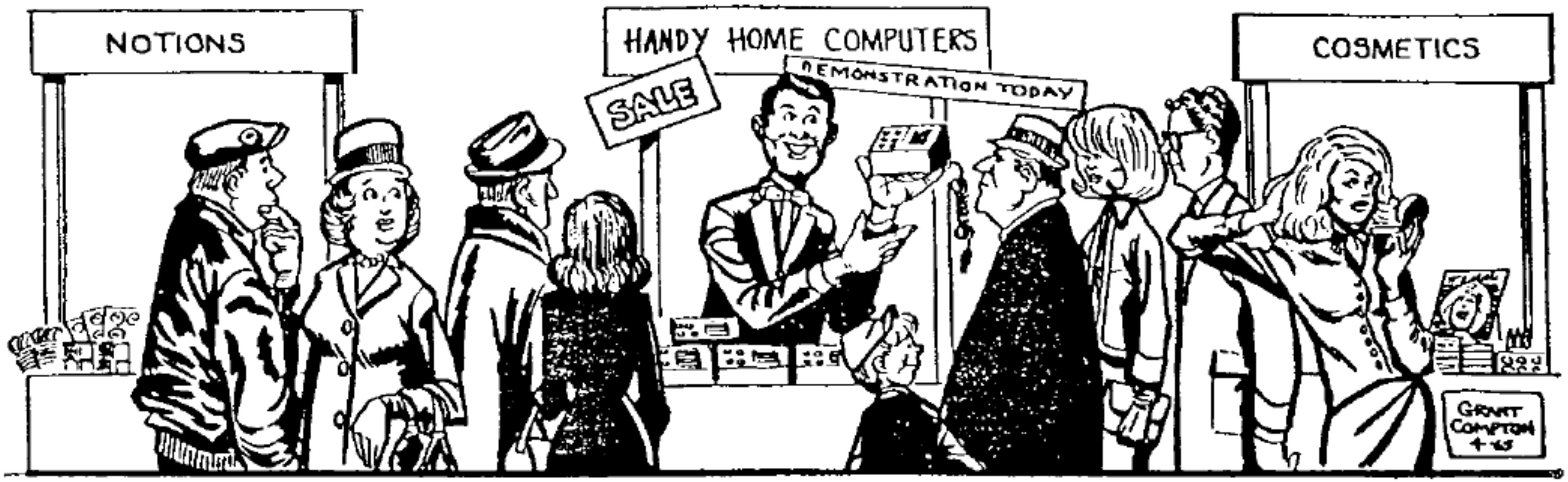
Predicted if speed of innovations in "cramming more components onto integrated circuits" kept up …



Moore's Law:

… we would soon be buying computers at the local market …

which apparently was a preposterous idea

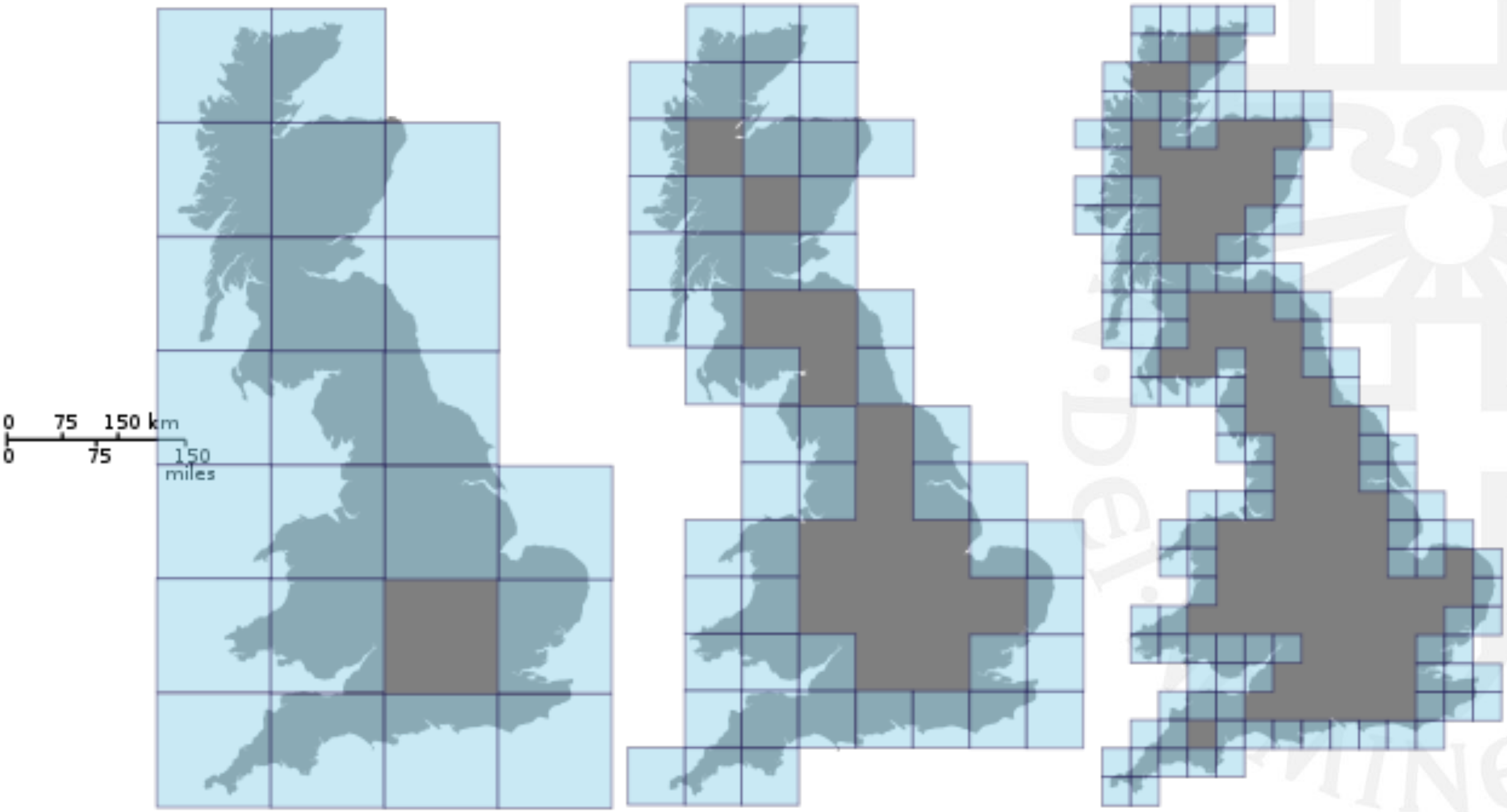


Moore, Gordon E. (1965). "Cramming more components onto integrated circuits"


VANAF WOENSDAG
26-02
ONZE AANBIEDINGEN
Tablet PC cover
6.99*
GSM AT-B26D
FM-RADIO
BLUETOOTH 2.1
OP=OP
Per stuk
49.99*
2 Jaar GARANTIE
10" Tablet MD98516
16 GB GEHEUGEN
Per stuk
179.00*
3 Jaar GARANTIE
ALDI TALK
€19.99
750 MB
€5
200 minuten
Wireless speaker adapter
14.99*
Bevestigingsset
2.99*
Hortensia
4.79*
ALDI
Aperitiefbiscuits
0.99*
Basis voor soep
1.29*
HONING DROP
MUNTEN DROP
Munt- of honingdrop
0.99*
Spijsbroodjes of appelflappen
1.19*
Mini-stroopwafels
1.89*
Smulgerechten
Chips patatje kapsalon of hete kip
0.99*
PDF
6-7/16
Index
Trefwoord


Kruidvat
voordeel magazine
Goed gekleed
DAMESJASJE
M t/m XL
39.99 19.99
true SPIRIT
Perfect haar
NIVEA HAARVERZORGING OF STYLING
1+1 GRATIS
Stralende lippen
MAX FACTOR LIPSTICK, -GLOSS OF -LINER
ALLES VOOR € 10.00 PER STUK
Lekker schoon
ROBOTSTOFZUIGER
Dirt Devil
129.99 79.99
STEEDS VERRASSEND, ALTIJD VOORDELIG!
Geldig van dinsdag 25 februari t/m zondag 9 maart 2014

How to describe scaling relations: Calculate a “fractional” dimension, e.g. box-counting dimension



Scaling phenomena

Measuring dimension

What is the dimension of these objects?:

Definition (Euclidian) dimension: *The number of degrees of freedom you have to move through a space.*

Definition of a space: *a collection of points*

0 1 2 3

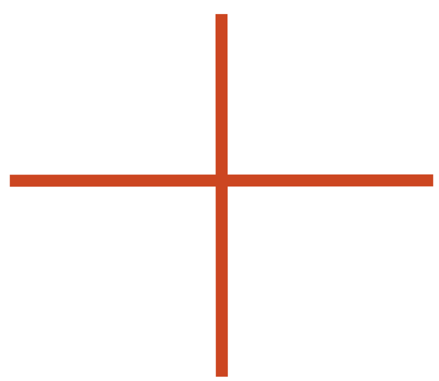
Scaling phenomena

Measuring dimension

What is the dimension of these objects?:

Definition (Euclidian) dimension: *The number of degrees of freedom you have to move through a space.*

Definition of a space: *a collection of points*



Locally 1 dimensional
On the cross-section 2 dimensional

Which one is it, the smallest?

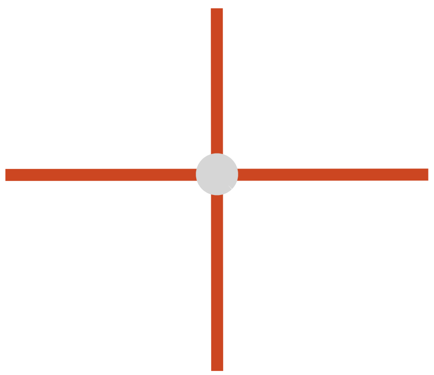
Then this space would be 0 dimensionaal, because a point = 0

Scaling phenomena

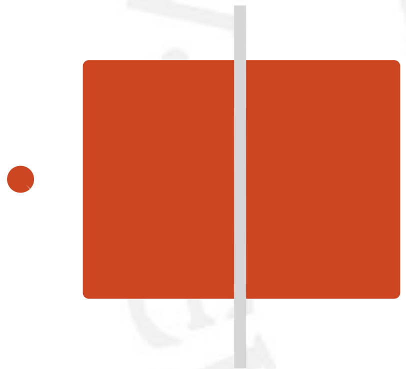
Measuring dimension

Topological Dimension

Introduces *local* and *global* dimension.
Global: Take the dimension of the object with which you can divide the space in two parts and add one.
Constraint: **The dimension must stay the same over linear transformations**



Point = 0 (+ 1)
1 dimensional



Local 0 and 2 dimensional
Take the highest
Global = 2 dimensional

Scaling phenomena

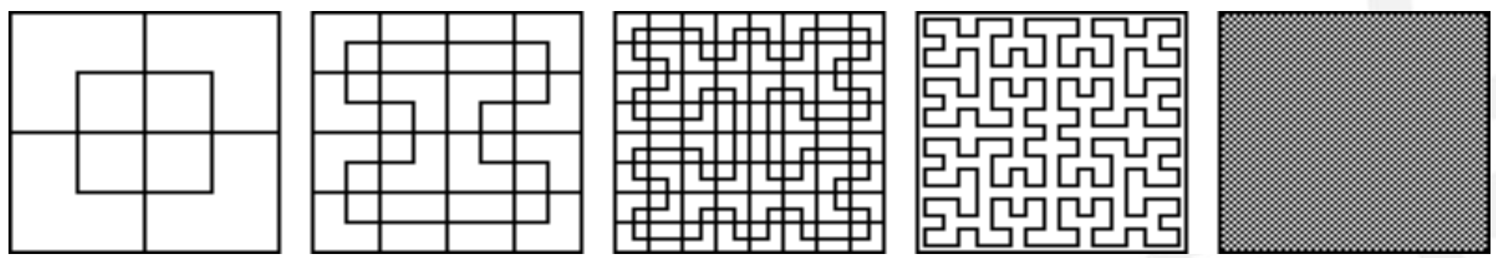
Measuring dimension

Constraint: **The dimension must stay the same over linear transformations**

Consider these linear transformations

Peano's Monster

line-like



plane-like

Koch's Coastline

line-like

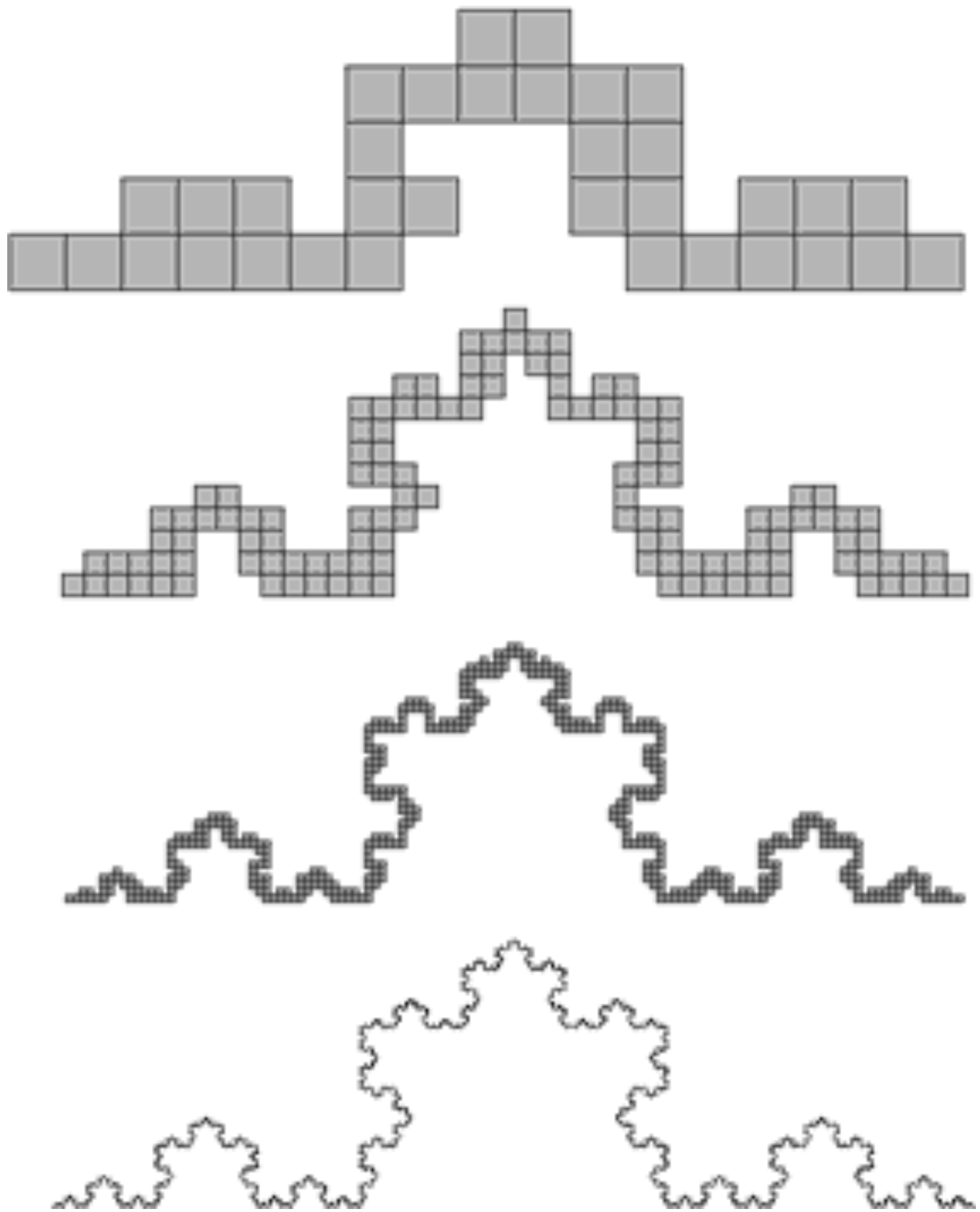
less line-like
(more complex)

Sierpinski's Gasket

plane-like

less plane-like
(more complex)

Solution:
Calculate a "fractional" dimension, e.g. box-counting dimension



Hausdorff-Besicovitch dimension

(box-counting dimension, covering dimension, packing dimension, mass-radius, circle-counting, etc, etc)

$$D = \frac{\log \mathbf{N_h}}{\log 1/\mathbf{h}}$$

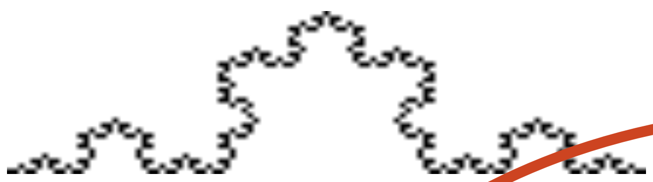
N= number of blocks of size **h** needed to cover the object

Relation between *measure stick* and *measurement outcome*, or: *"scaling of bulk with size"*

Scaling phenomena

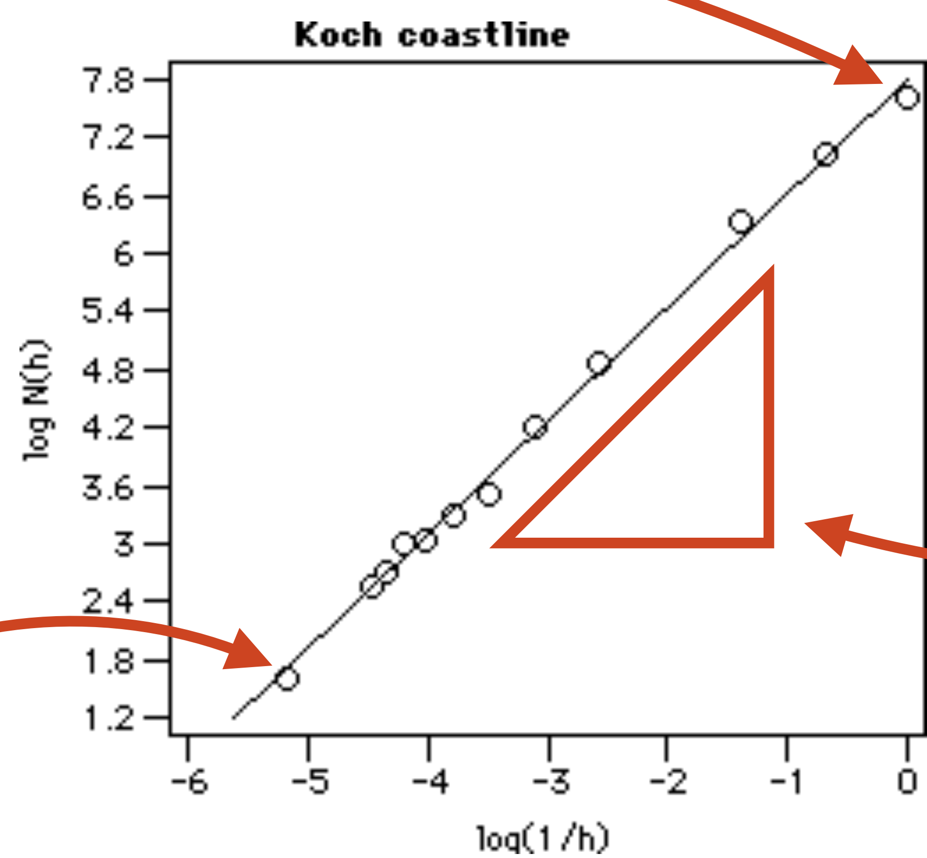
Measuring dimension

Koch Coastline



log (1/h)	log N(h)
0.00	7.61
-0.69	7.04
-1.39	6.33
-2.56	4.86
-3.09	4.22
-3.50	3.53
-3.78	3.30
-4.01	3.04
-4.19	3.00
-4.34	2.71
-4.48	2.56
-5.18	1.61

dimension (experimental) = **1.18**
dimension (analytical) = **1.26**
deviation = 6%

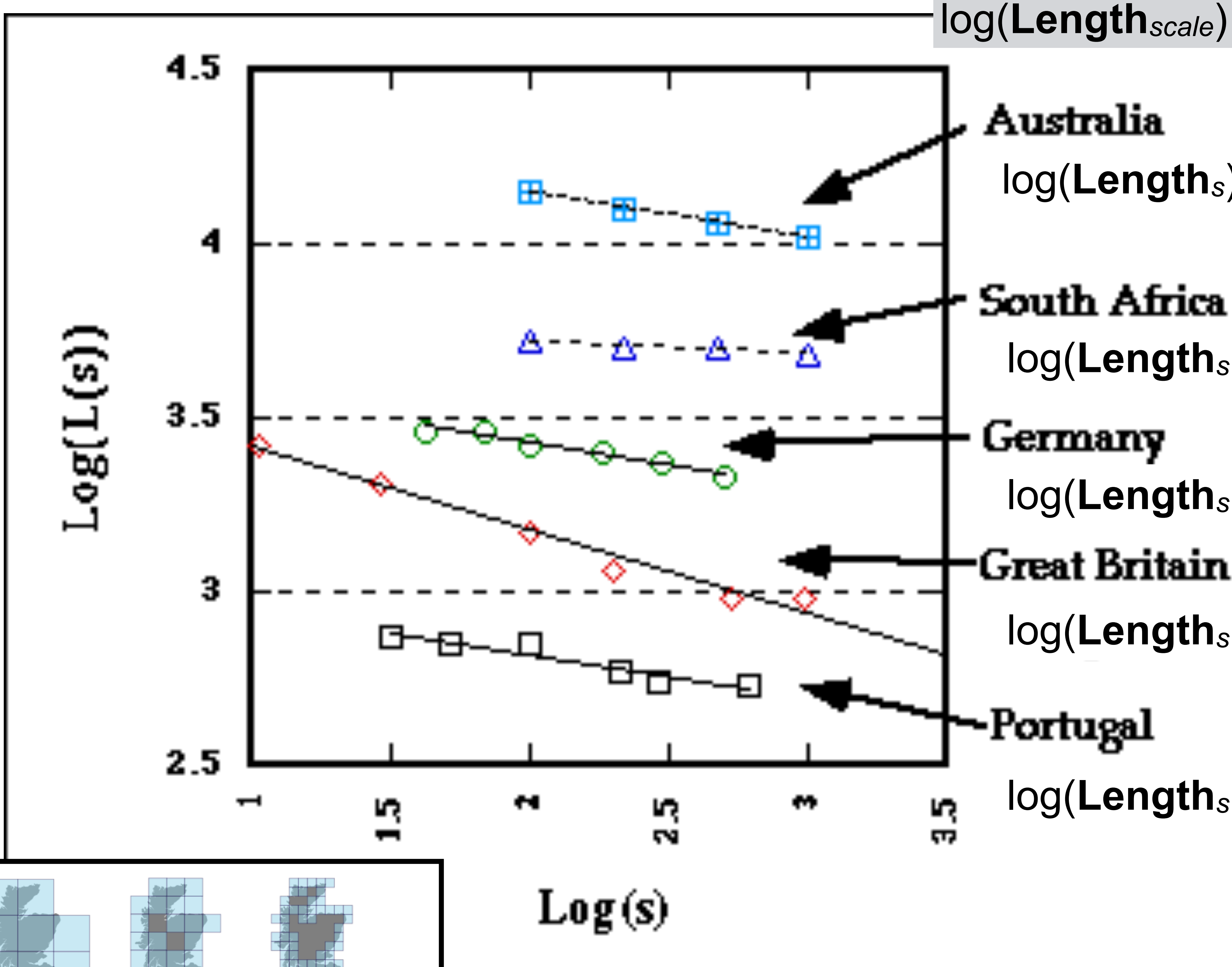


Fractal dimension
it's a fraction!

```
Call:
lm(formula = L ~ invS, data = df)

Residuals:
     Min       1Q   Median       3Q      Max
-0.18777 -0.06292  0.02390  0.06059  0.16703

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.79777    0.07318  106.55  < 2e-16 ***
invS         1.17611    0.02109   55.75 8.35e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1137 on 10 degrees of freedom
Multiple R-squared:  0.9968,	Adjusted R-squared:  0.9965
F-statistic:  3109 on 1 and 10 DF,  p-value: 8.355e-14
```

Scaling phenomena

$\log(\textbf{Length}_{scale}) = b*\log(scale)+a$



Australia: $\log(\textbf{Length}_s) = -.13*\log(s)+4.4$

South Africa: $\log(\textbf{Length}_s) = \mathbf{-.03}*\log(s)+3.8$

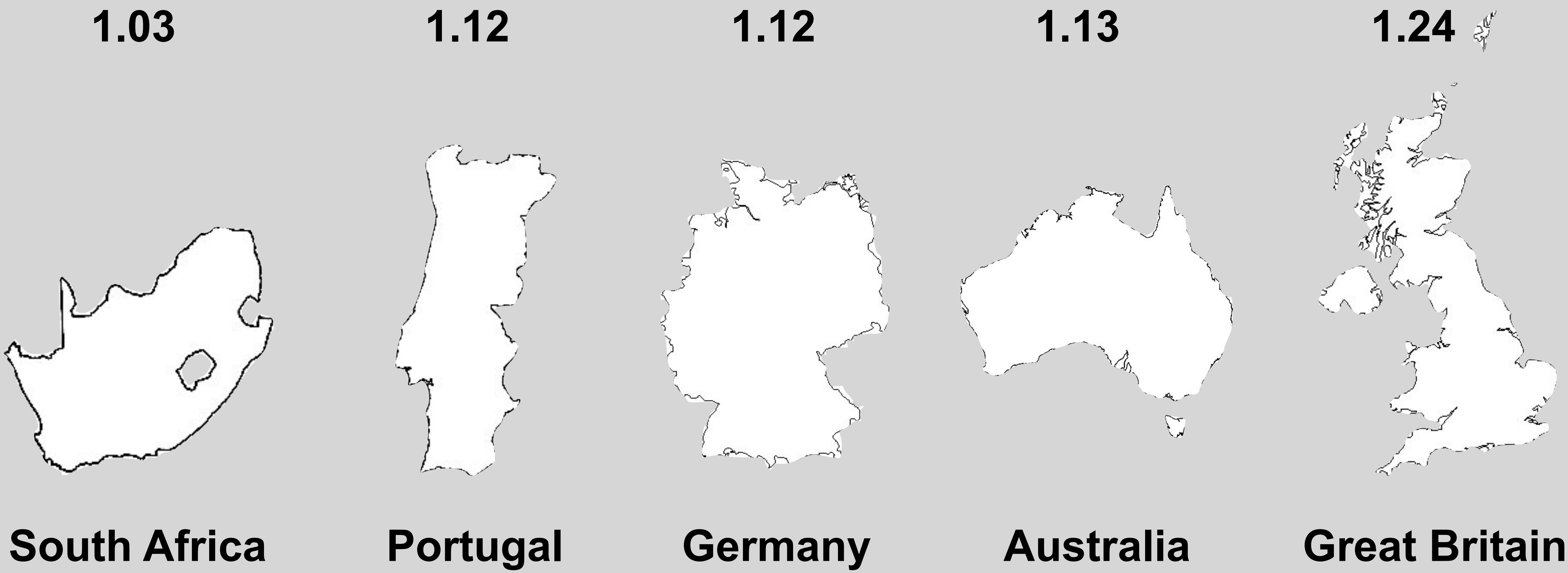
Germany: $\log(\textbf{Length}_s) = -.12*\log(s)+3.7$

Great Britain: $\log(\textbf{Length}_s) = \mathbf{-.24}*\log(s)+3.4$

Portugal: $\log(\textbf{Length}_s) = -.12*\log(s)+3.3$

Scaling phenomena

Scaling and Complexity



Ordered by scaling exponent, the log-log slope