3b. Re-cap: Valuation & verification

Bringing together...

Basic checks

- Range and limits
  - Realistic values
  - Beware of bias (e.g. the 'ozone hole', 1986)
- Logical consistencies
  - max(x) > min(x), etc
  - Inversion
- Maths
  - e.g. Benford's law (“the first digit law”), binominal distribution, Poisson distribution. ...

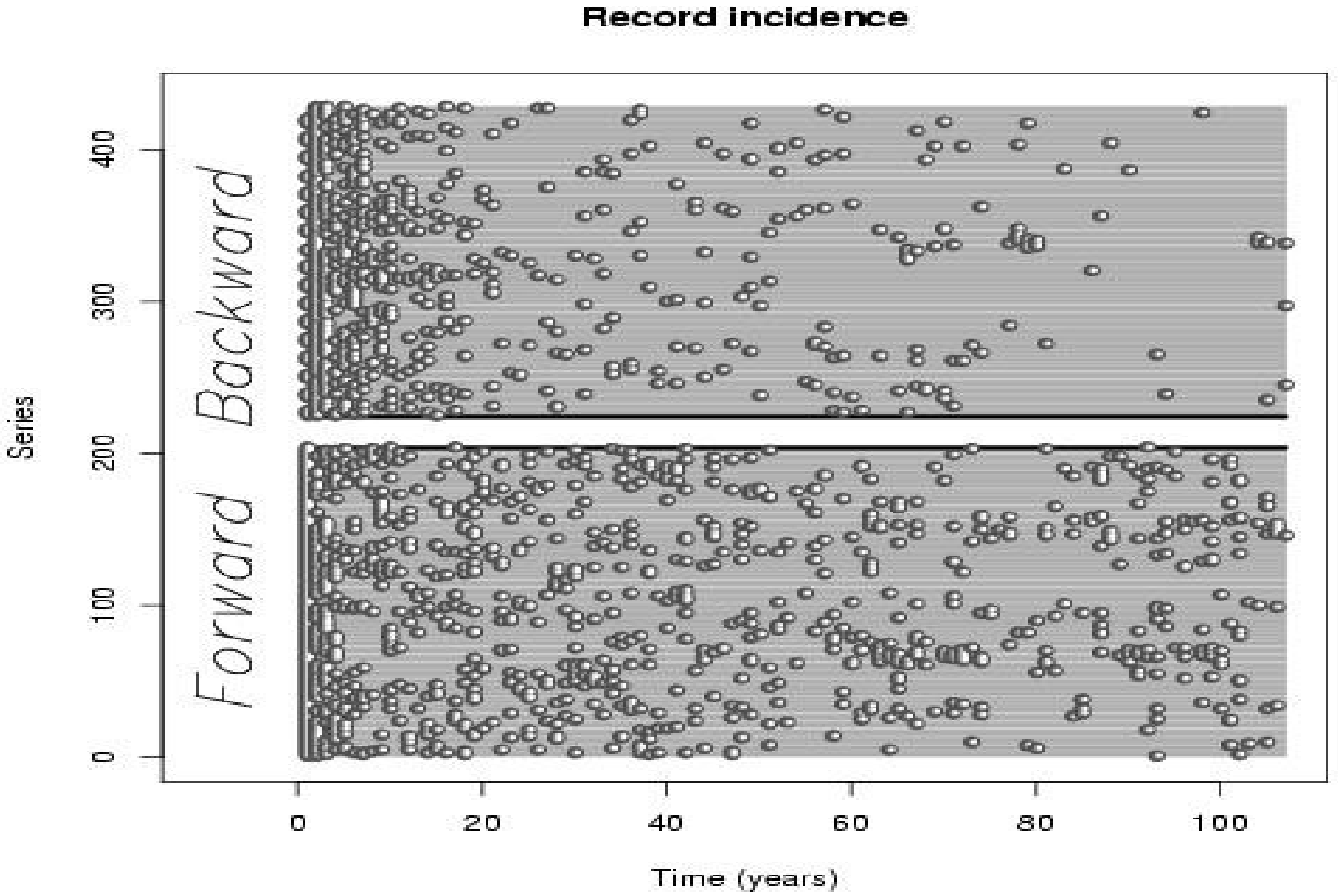
$$P(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}\left(1 + \frac{1}{d}\right).$$

V&V: is the data iid?

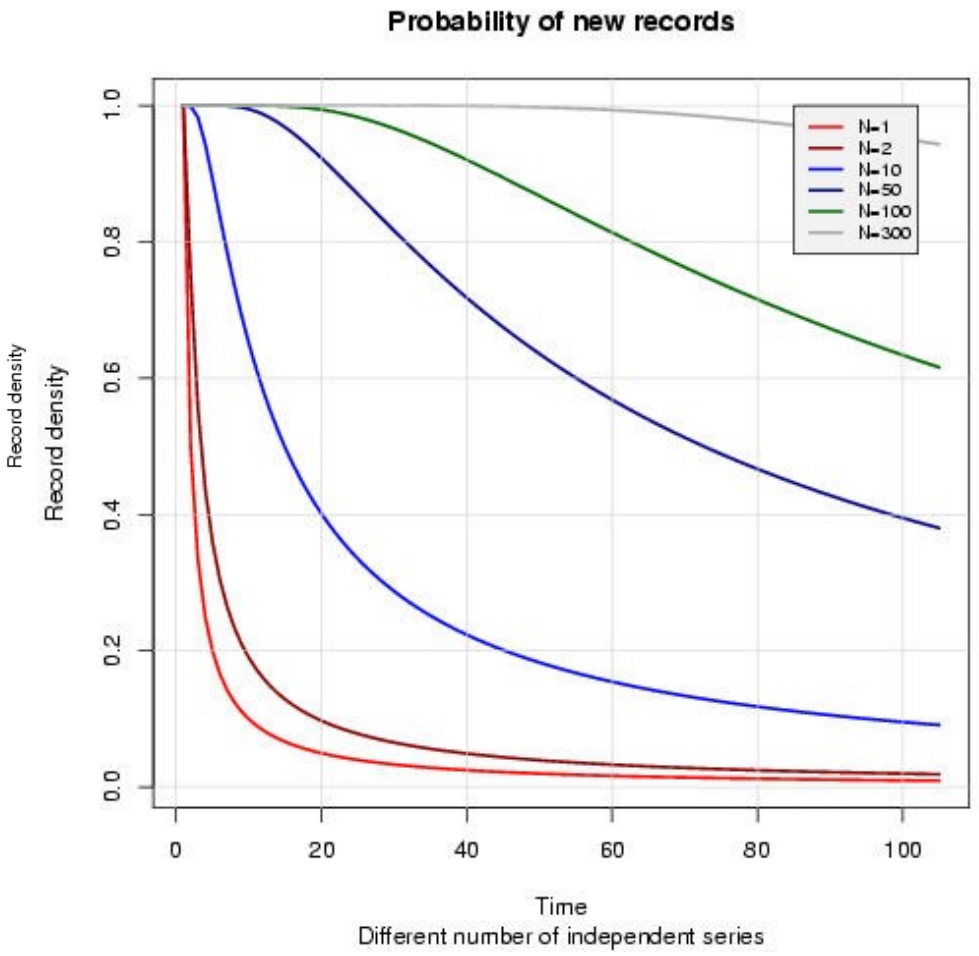
- **I**dentically and **i**ndependently **d**istributed?
- Does the pdf change over time? (data not identically distributed).
- *Pr(x = max(X)) = 1/n*
- V&V for **methods assuming stationarity** (return value analysis, analogs)

Number of record-breaking events

$x(t) > x(t')$ for all $t' < t$



Parallel series – *Pr(at least 1 record-breaking event)* (binomial distribution)

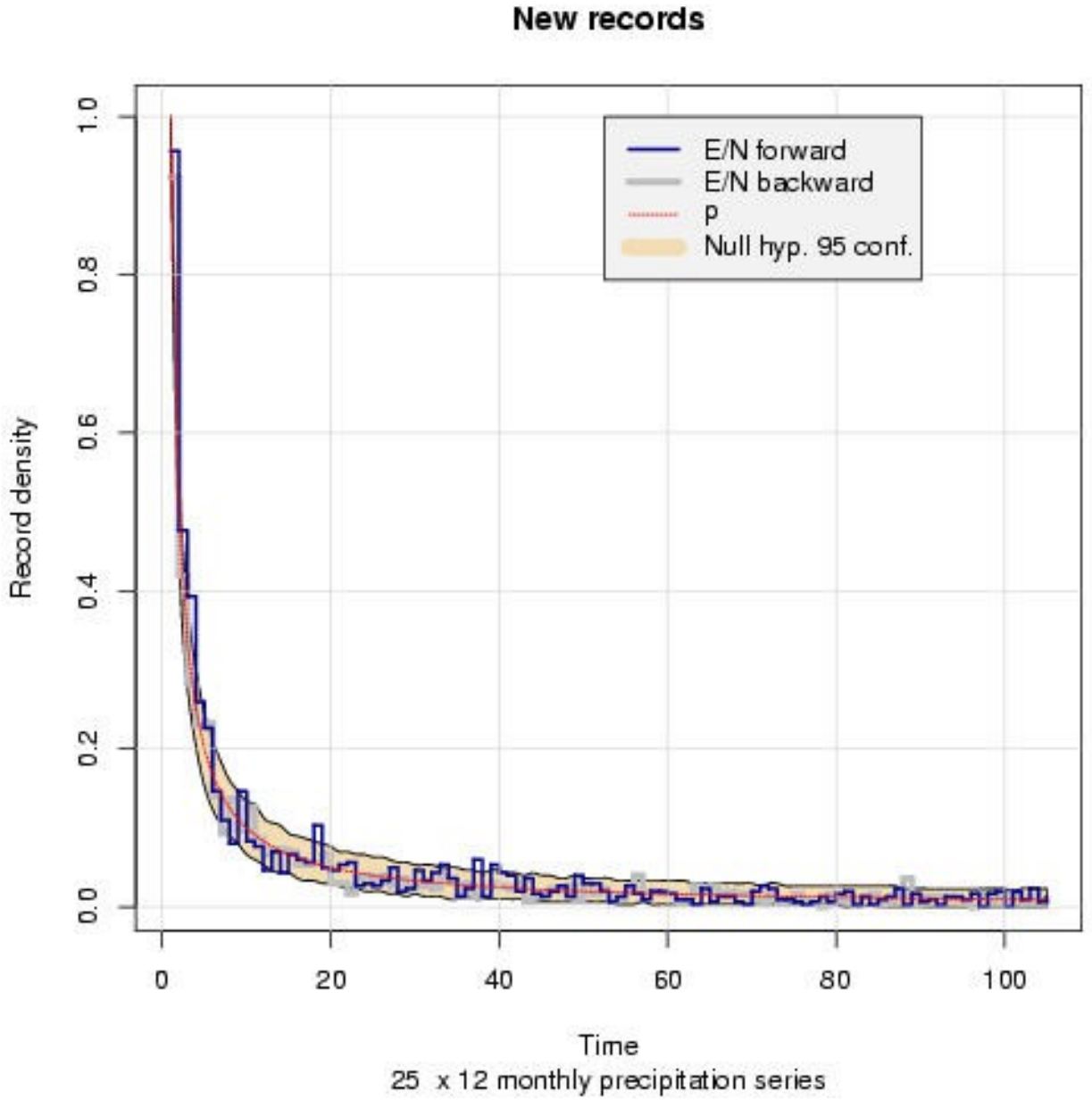


A random variable of rational numbers with independent and identical distribution (iid) has following property:

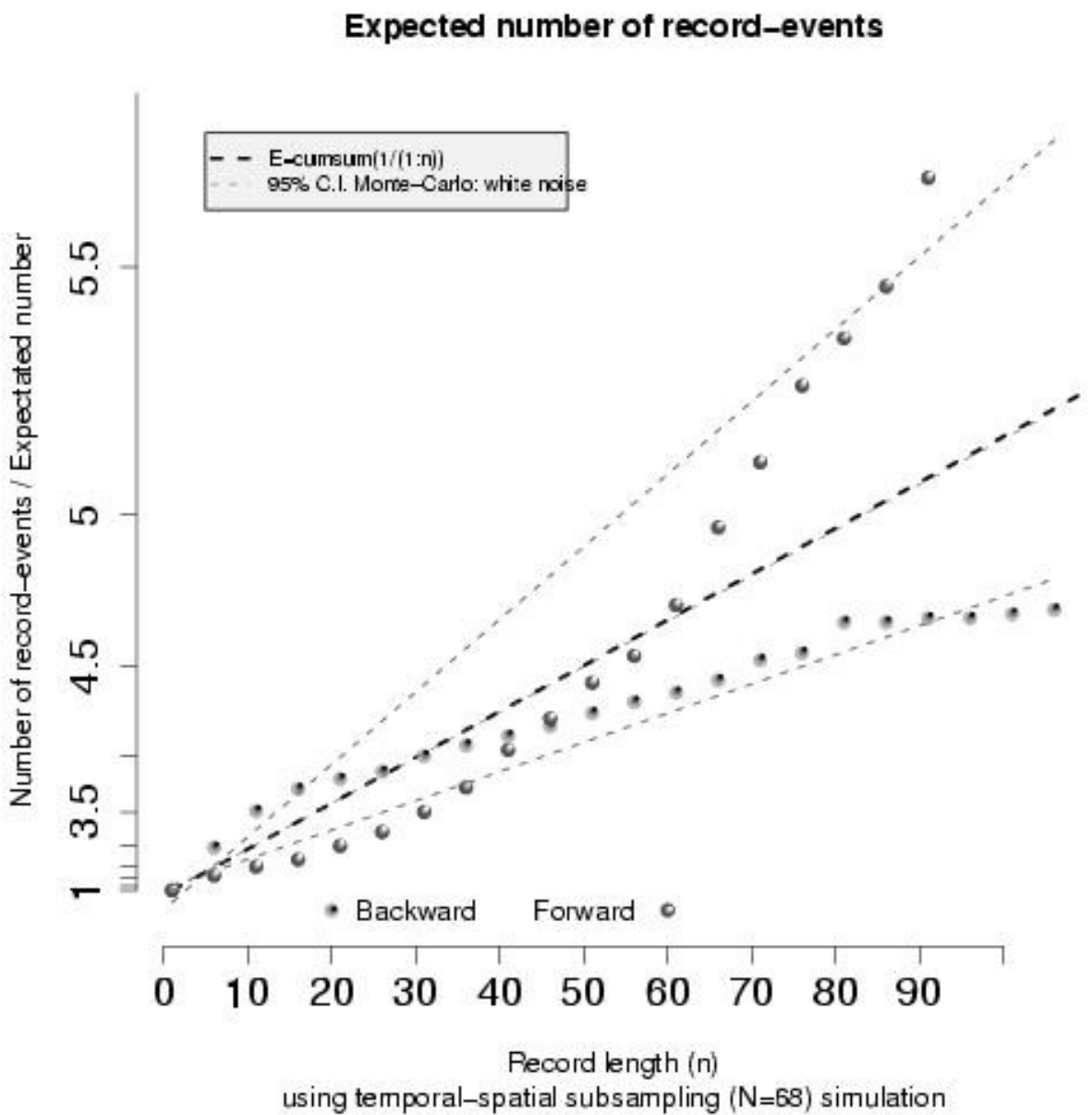
Pr(n=record) = 1/n

(assuming no ties)

Record-event statistics

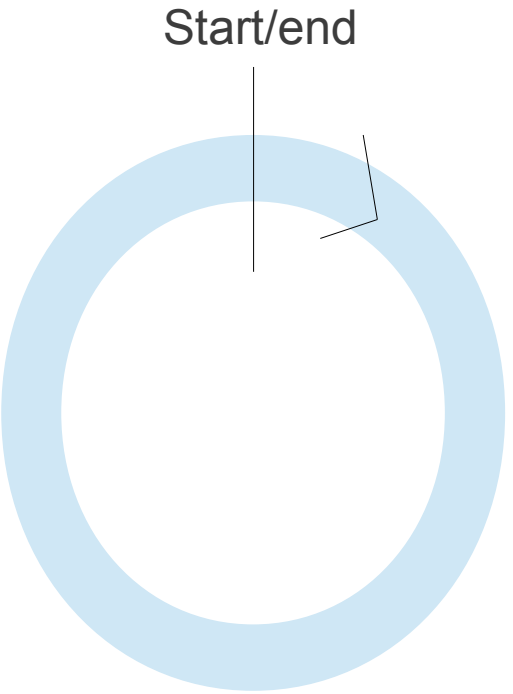


The iid-test.


Expected number of record–events
E-cumsum(1/(1:n))
95% C.I. Monte–Carlo: white noise
Number of record–events / Expectated number
1
3.5
4.5
5
5.5
Backward
Forward
0
10
20
30
40
50
60
70
80
90
Record length (n)
using temporal–spatial subsampling (N=68) simulation

Checks & tests

- Sample tests
- Special cases
- Consistency
  - Inversion/reversion
  - Full-circle
- Monte-Carlo simulations



Example: Full circle test...

- **Google translate**: English → German → Norwegian → English:
- **Input**: “Validation and verification can be implemented through a circular path where the end-result is the original input”
- **Output**: “Validation and verification can be performed through a circular path where the end result is the original input”

Strengths of computers

- Fast,... many iterations
- Random number generators
- Constructed positives.
- Wide range of inputs – robustness.
- Creative solutions.

Monte-Carlo simulations

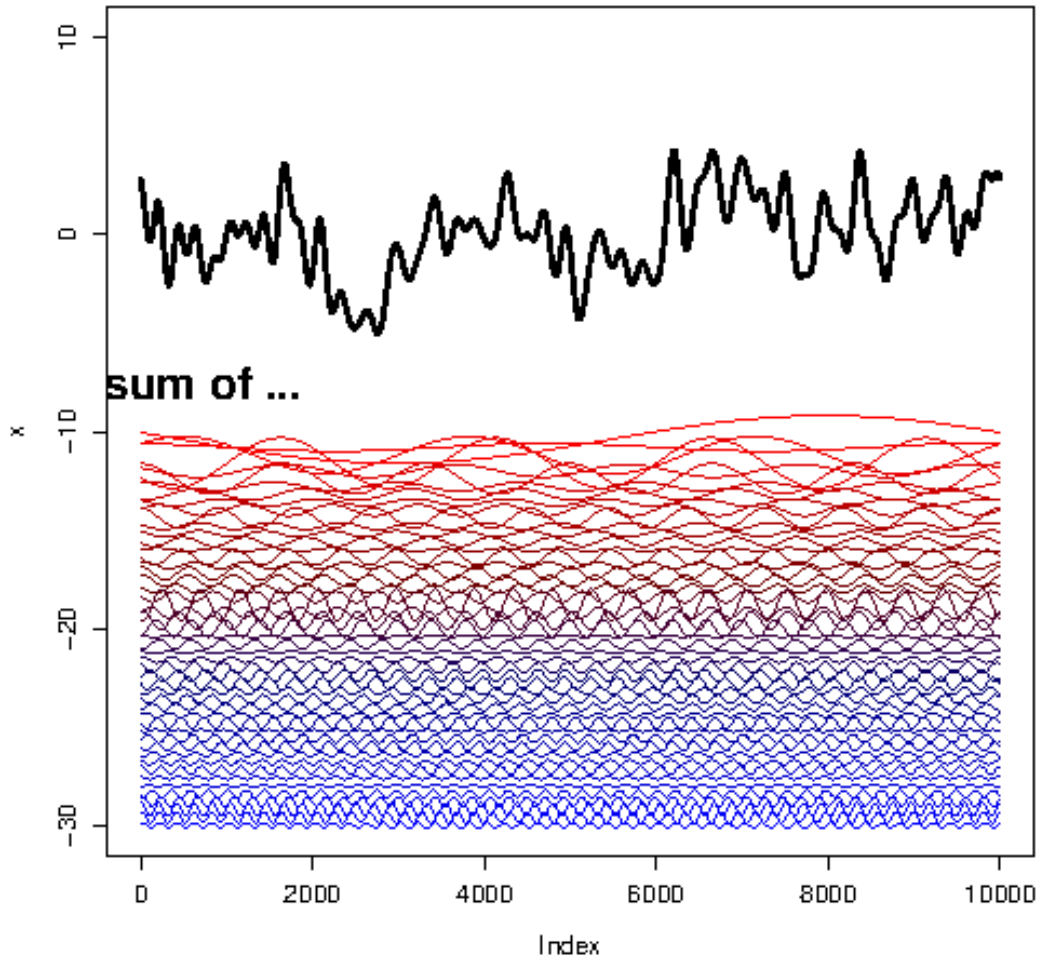
The use of random number generators to test methods and hypotheses. A priori knowledge: randomness – no skill. Digital version of rolling dice.... and roll the dice very many times...

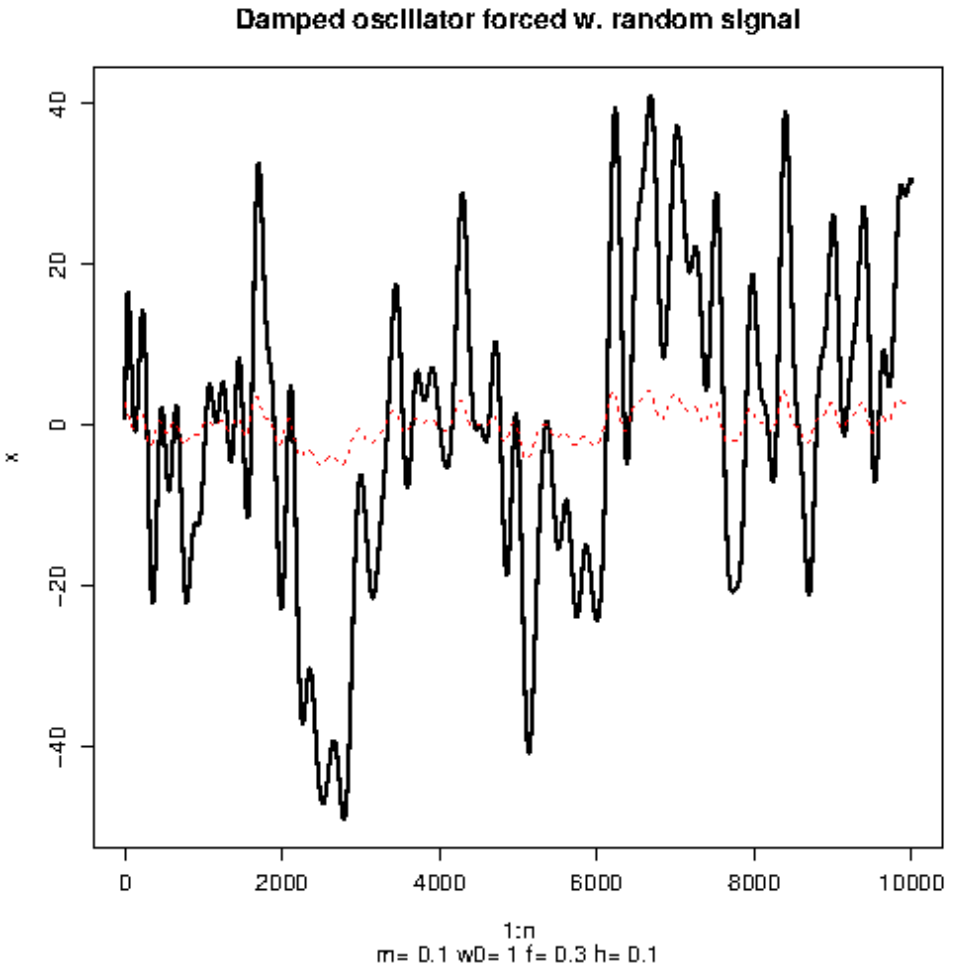
Monte-Carlo simulations...

- Test the null hypothesis
- See if we can get the same results by chance
- Mimic the process without the dependency
- Null-hypothesis: what you can get with noise.
- In addition, input can be data with a given positive result.

Code development – additional tests

- Simulate 'noise' – lack of signal/dependency





Example: CCA

- **C**anonical **c**orrelation **a**nalysis (CCA).
- Identify pairs of patterns with the greatest correlation.
- V&V using Monte-Carlo techniques to sample many different combinations

Example: CCA

- Canonical correlation analysis (CCA) → patterns with the highest correlation.

Two data sets

$$Y = \mathcal{G}\mathcal{U}^T,$$
$$X = \mathcal{H}\mathcal{V}^T,$$

Covariance matrices:

$$C_{YY} = YY^T,$$
$$C_{XX} = XX^T,$$
$$C_{YX} = YX^T.$$

$$C = C_{YY}^{-0.5} C_{YX} C_{XX}^{-0.5},$$

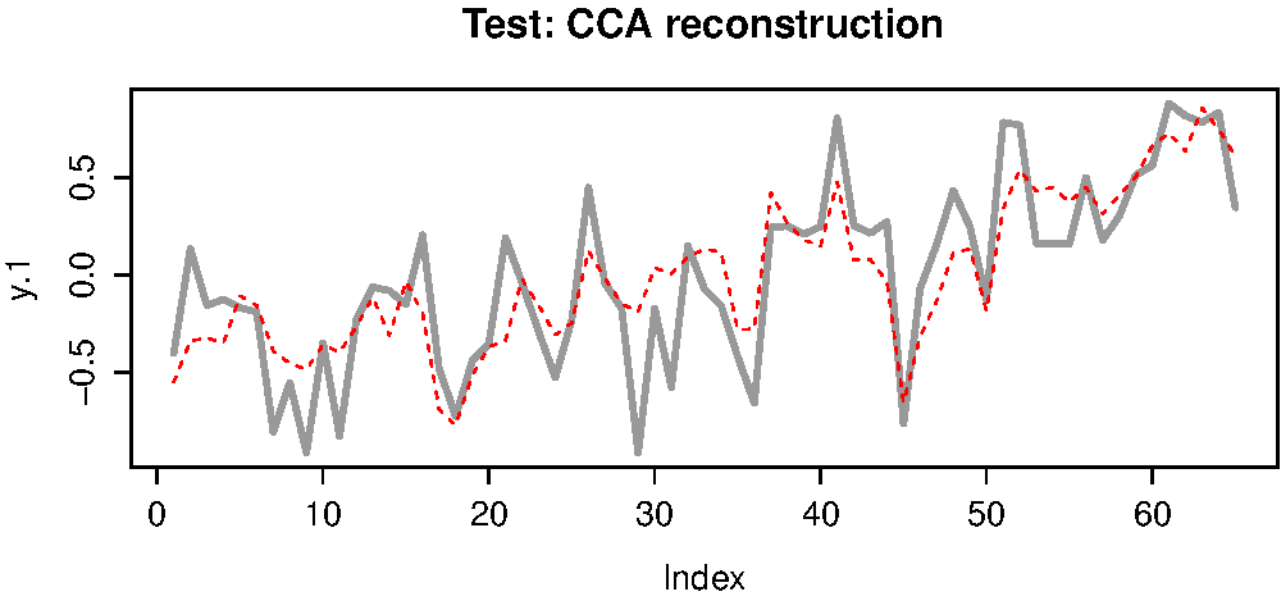
Eigenvalue problem (SVD)

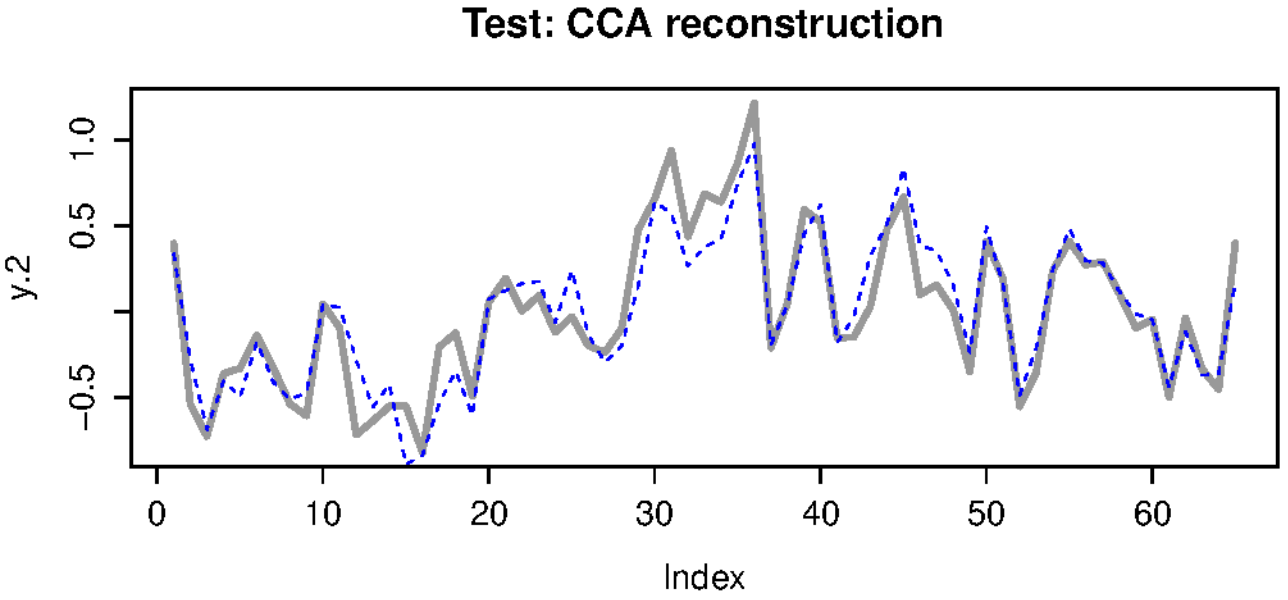
$$C = LMR^T.$$

clim.pact: 'ABC4ESD()'

Invert the problem

Start from the other end and work backwards 'CCA(..., test=TRUE)' [clim.pact]

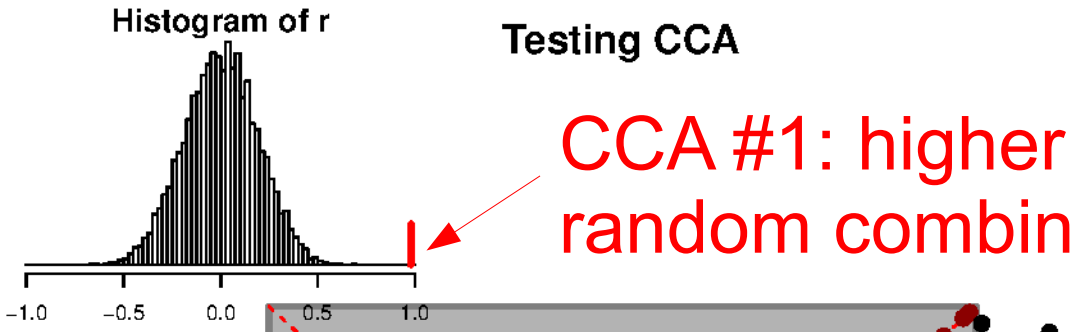




Other aspects with CCA

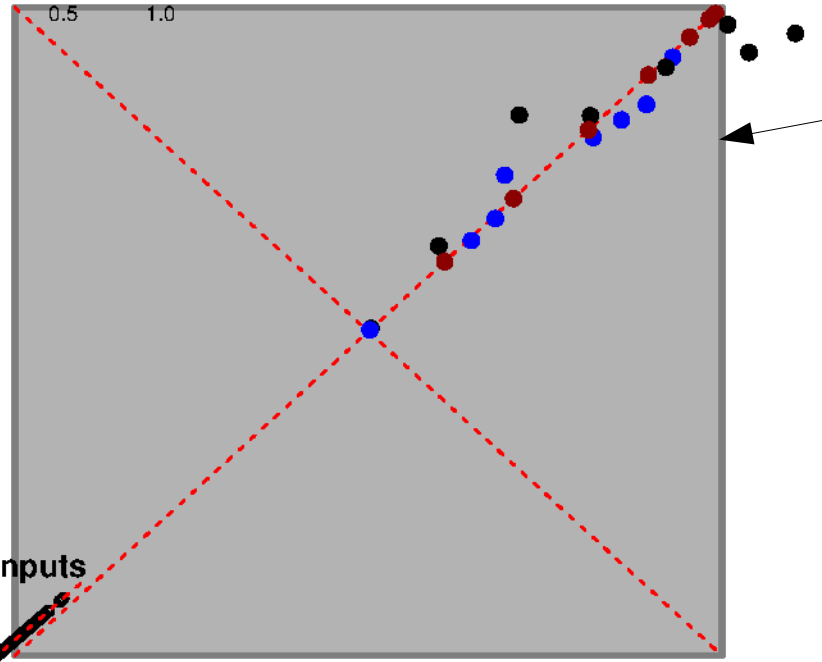
- Test the limit
  - CCA is supposed to find patterns which have the greatest possible correlation in terms of the information embedded in the data.
  - Can we find cases with higher correlation?

Examine the limits...

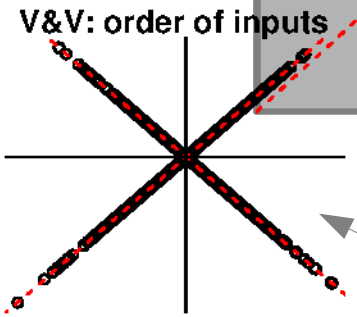


Testing CCA

CCA #1: higher than most random combinations



Correlation between series & from matrix *M*: Fairly good match.



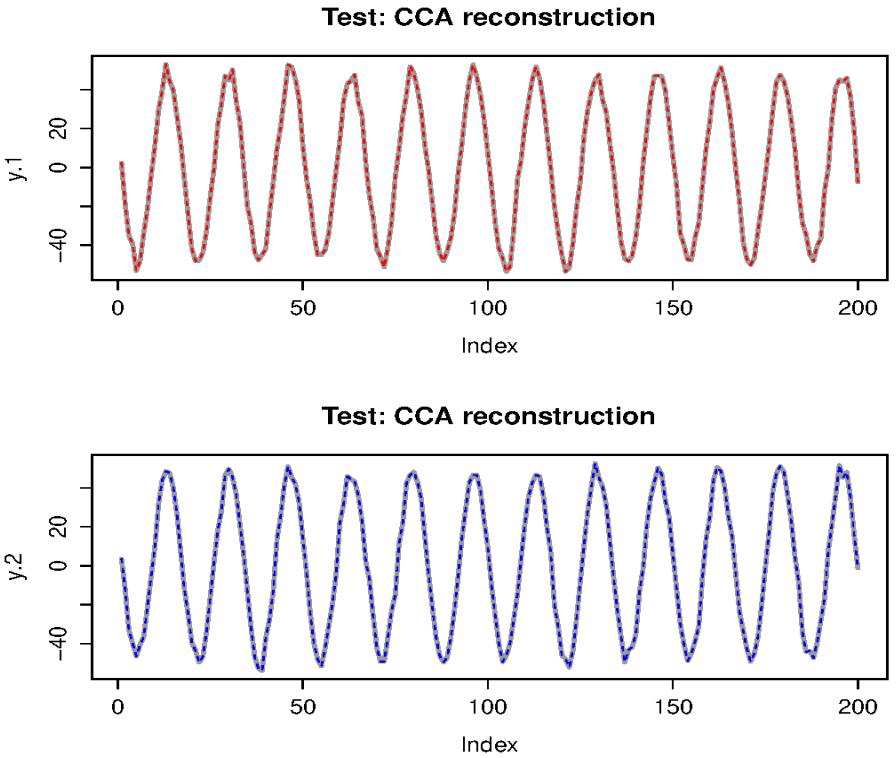
Order: CCA(X,Y)=CCA(Y,X)

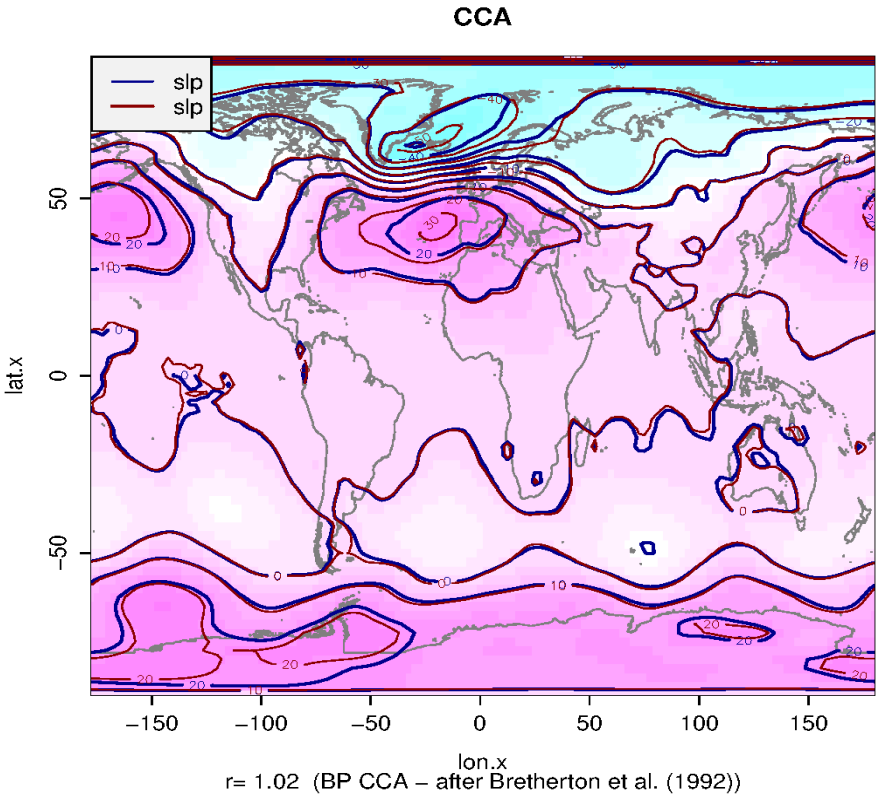
est. cor

r from SVD

Insert known signal

'testCCA()' [clim.pact]





Regression

- Analysis of variance (ANOVA)
- P-values, variance (R^2), F-ratio, etc.
- Ordinary linear regression (OLR), generalised linear models (GLM; maximum likelihood), factorial regression.
- Univariate, multiple & multivariate.
- Care – easy to get astray.

Challenges for V&V

- Time scales for archiving digital information
- System for organising work
  - In-line comments
  - Documentation
  - Testing
- Access & transparency
- Agnotology
- Bias & perception
- Unexpected situations, new data (often a bug)

The CRAN repository

- Never guaranteed a bug-free code
- Sharing methods and data means more testing.
- Building R packages, involves a numper of testing procedures
  - `R CMD check package.folder`
  - `Structure & documentation`
- Many users with different data, situations, etc – feedbacks.

Community

- Discussions is an inherent part of science
- Differences are opportunity for progress
  - This is not always perceived like this
  - Defencive attitudes
- Trial and failure are essential concepts and valid strategy
  - Sometimes a scientist needs to be brave
- Social aspects can be difficult
  - 'Office politics', competition, envy, ...

Active Internet

- RealClimate.org
- Comment & discussion
- The value of social input and different views
- Accessibility
- New ideas → publications.

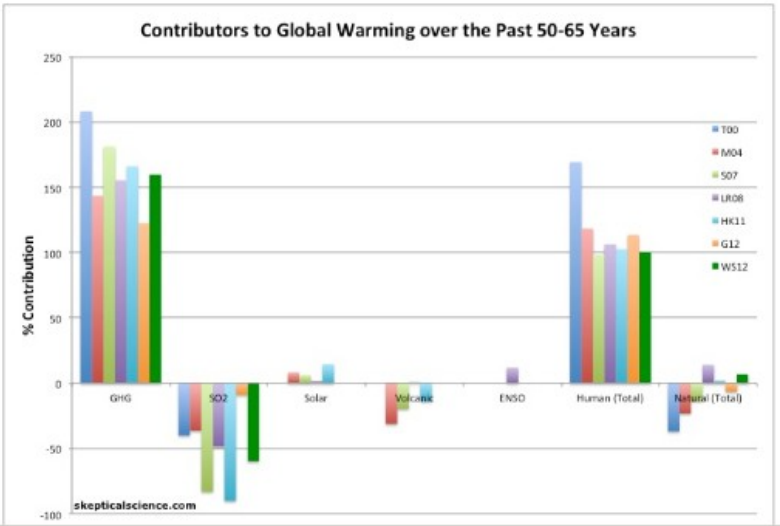
IPCC Draft Report Leaked, Shows Global Warming is NOT Due to the Sun

Posted on 14 December 2012 by dana1981

This post has been re-published by *The Guardian*

Alec Rawls, an occasional guest poster on the climate contrarian blog WattsUpWithThat who signed up to review the upcoming Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (as anyone can), has "leaked" a draft version of the report and declared that it "*contains game-changing admission of enhanced solar forcing*." This assertion was then repeated by James Delingpole at *The Telegraph* (with some added colorful language), and probably on many other climate contrarian blogs.

If the IPCC was to report that the sun is a significant player in the current rapid global warming, that would indeed be major news, because the body of peer-reviewed scientific literature and data clearly show that the sun has made little if any contribution to the observed global warming over the past 50+ years (Figure 1).



Propagating knowledge

- Persuading your colleagues
- Sanity control
- Publication – the 'gold standard'
  - Merit & prestige
  - Establish community knowledge
  - Benefit for society
  - Funding & utility
- Acessability: visibility + legitimacy

Keep in mind

- **Falsification!** Never 100% verification
  - Typical: code fails with new data.
  - Unforeseen situations/possibilities
- Bias and deceptive results – expectations
- Dimensional analysis
- Monte-Carlo simulations & test input
- Inversion, reverse, full-circle consistency
- Documentation

Final word

- Practising V&V may open eyes to more . It may be like a curse, as the world may prefer to live with delusions.

V&V Resources

- Further reading:
  - Jolliffe, Ian T., and David B. Stephenson, eds. Forecast Verification: A Practitioner's Guide in Atmospheric Science. Wiley, 2003.
  - Magee, B. Popper. 13th ed. Fontana Modern Masters. London, Britain: Fontana Press, 1973.
  - A. Engel, Verification, Validation, and Testing of Engineered Systems (Wiley Series in Systems Engineering and Management), 2011
- URLs.
  - Facebook page: agnotology

+ references provided on the previous slides