Misuses of Statistical Analysis in Climate Research

Hans von Storch
Max-Planck-Institut für Meteorologie
Hamburg, Germany

Talk presented at the
6 IMSC, Galway
June 1995

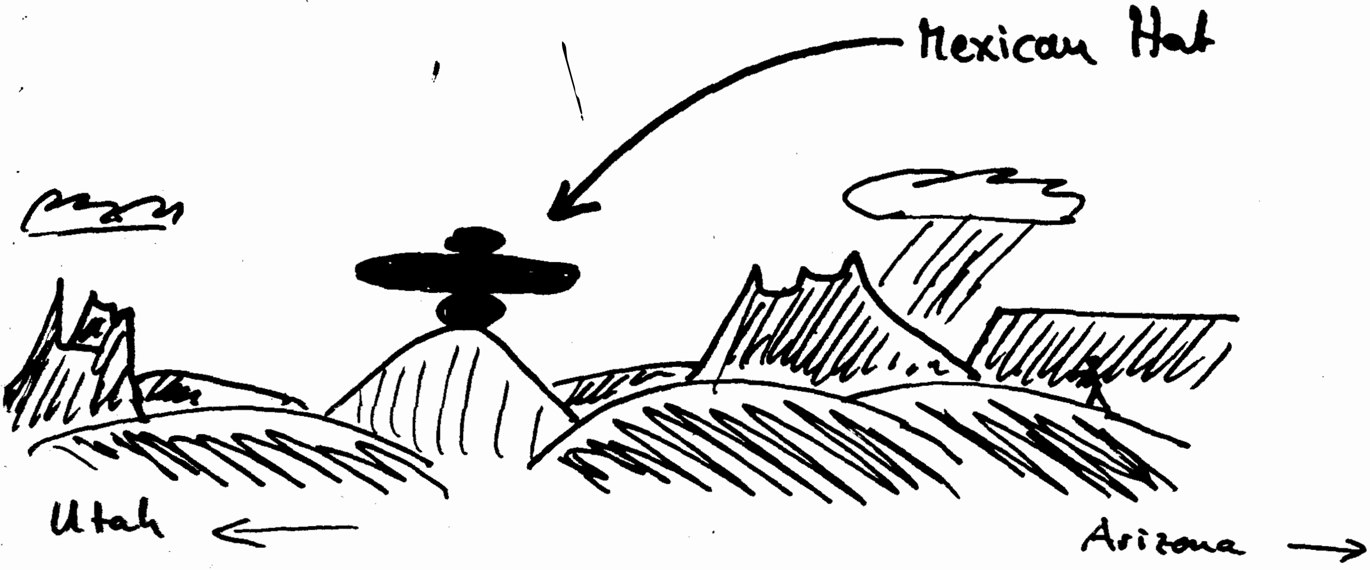
See detailed paper in
von Storch and Navarra (eds.):
Analysis of Climate Variability -
Applications of Statistical Techniques.
Springer Verlag 1995, 334 pp

Many Misuses Arise from ...

- **Obsession with statistical recipes** such as statistical testing.
- **Use of statistical techniques** in a cook book like manner.
- **Misunderstanding of names** such as the decorrelation time.
- **The faith in results obtained with sophisticated techniques.**

OBSESSION
with statistical testing

The MEXICAN HAT



H_0: Mexican Hat is natural

What is the probability to observe three stones that form a Mexican hat just by natural processes?

Walk thru the desert and sample. Check 10^6 triples of stones.

Result of survey:

mexican hats $= 1$

other combinations $= 10^{+6} - 1$.

$\Rightarrow p(y = \text{Mexican Hat} \,/\, y \text{ is natural}) \lessapprox 10^{-6}$.

$\Rightarrow$ Reject H_0, accept H_A = Mexican Hat is man-made.

But H_0 is natural.

?

Mexican Hat.

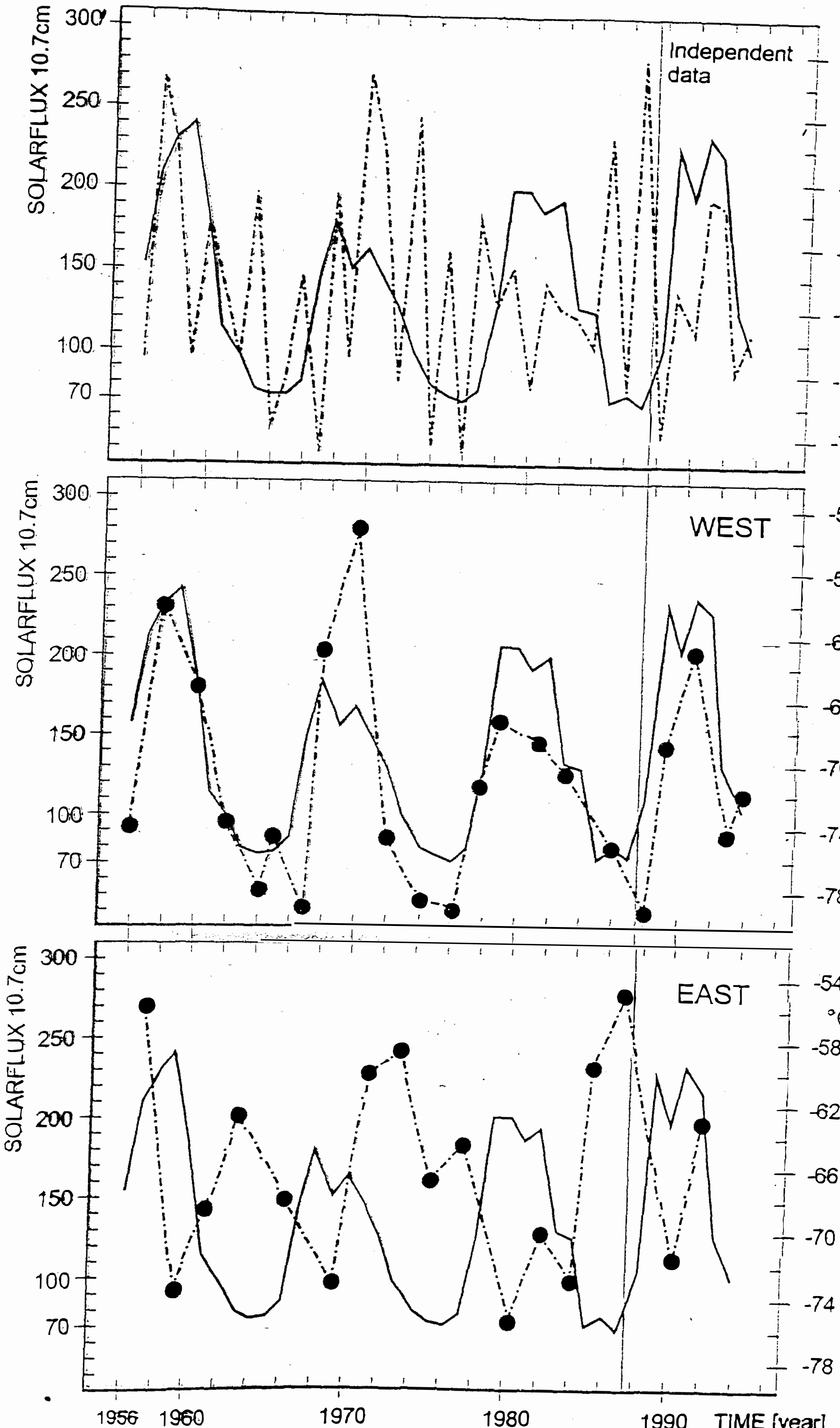
Problem is:

Nullhypothesis was formulated
<u>after</u> the observation has been made.

Example: Solar flux, the QBO and
stratospheric data.
(Labitzke & van Loon)

<u>Question</u>: Is any analysis of historical data
a case of a Mexican Hat?

North Pole 30 hPa JF temperature



J94HVSaa.ds4

Labitzke & van Loon, 1988

QBO in

← EAST or WEST phase →

The Case of
Misleading Names ...

... the **DECORRELATION TIME**

(see Thiébaux & Zwiers, 1984)

t-test continued

The EQUIVALENT SAMPLE SIZE

---

nullhypothesis $\mu = 0$.

n samples available to estimate $\hat{\mu}$ and $\hat{\sigma}$.

$$t = \frac{\hat{\mu}}{\sqrt{\frac{1}{n}} \cdot \hat{\sigma}}$$

$Var(\hat{\mu}) = \frac{1}{n} \cdot \sigma^2$ if the samples are independent.

$$\Rightarrow t \approx \frac{\hat{\mu}}{\sqrt{Var(\hat{\mu})}}$$

Equivalent Sample Size.

What if $\{x_1, \ldots x_n\}$ are serially correlated. Let's assume

$$x_{t+1} = \alpha x_t + \text{white noise}$$

Then the t-test is not applicable.

BUT $\quad \mathrm{Var}(\bar{x}_n) = \sqrt{\frac{1}{n_e}}\, \sigma^2 \quad$ with $\quad n_e = \frac{1-\alpha}{1+\alpha}\, n$

so that

$$t_e(x_1, \ldots x_n) := \frac{\bar{x}_n}{\sqrt{\frac{1}{n_e}} \cdot \sigma} = \frac{y}{\sqrt{\mathrm{Var}\, y}}$$

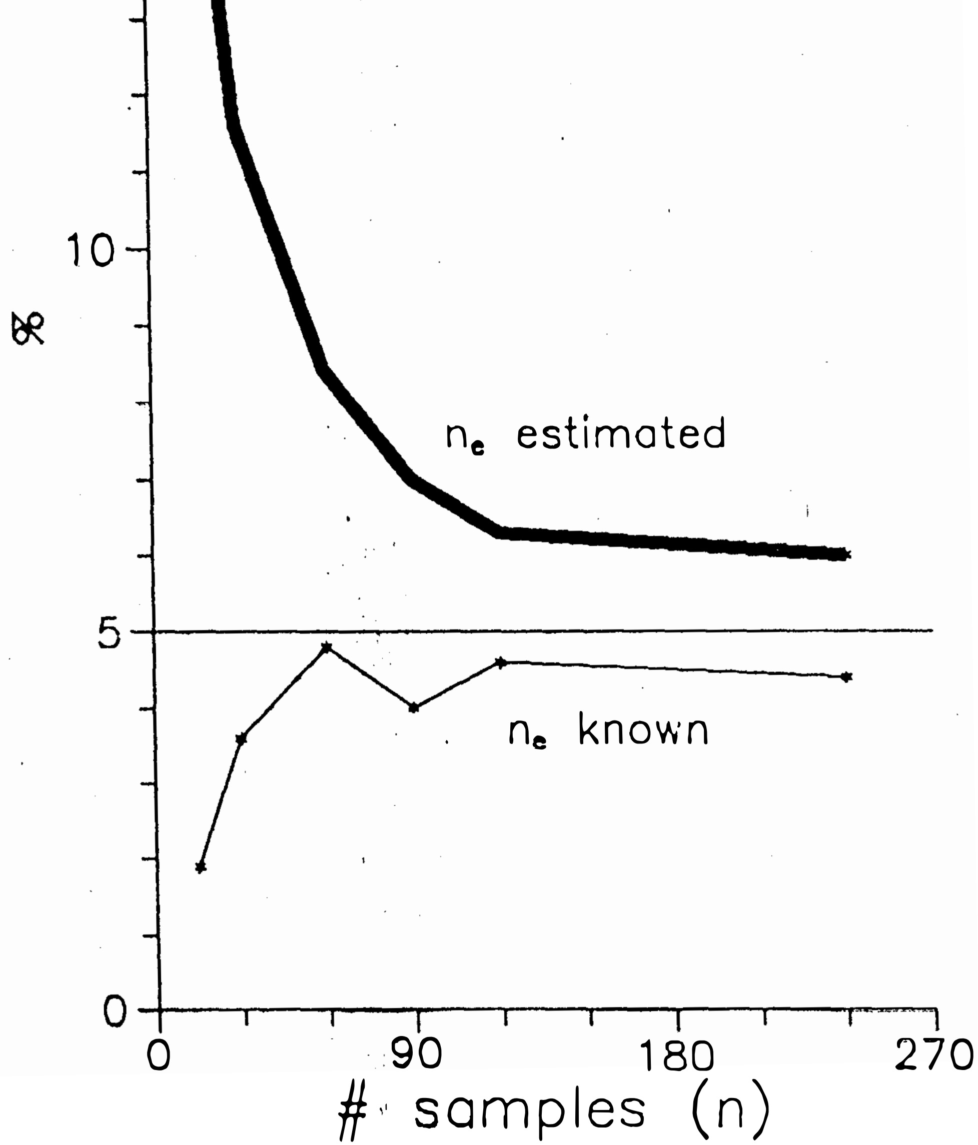
⇒ RECIPE

1) Calculate / estimate n_e from data
2) Calculate t_e
3) Calculate p-value of t_e from t-distribution with $n_e - 1$ degrees of freedom
4) Make decision on nullhypothesis

Fine ?

... We made a Monte Carlo experiment with an AR(1)-process, $\alpha = 0.6$ and variable n. With correct n_e and estimated n_e.

Rate of rejections of the nullhypothesis obtained with a conventional t-test utilizing the concept of the number of efficiently independent samples n_e to account for the serial corellation.



100 Monte Carlo Simulation were done for an AR(1)-process with memory 0.6.

NOTEBOOK C: GO GRA DAT N→E
11 JAN 1993 HvS

What to do?

... if $n_e \geq 30$ use Gauß-test

.... if $n_e \leq 30$ use "Table-look-up" test
(Zwiers & von Storch, J. Climate 1995)

Decorrelation time τ of an AR(1)-process

$$(*) \qquad X_t = \alpha X_{t-1} + Z_t$$

[with $\alpha < 1$, Z_t = white noise] is

$$\tau = \frac{1+\alpha}{1-\alpha} \geq 1$$

Process $(*)$ is equivalent to

$$(**) \qquad X_{t+k} = \alpha^k X_{t-1} + Z'_t$$

with white noise Z'_t.

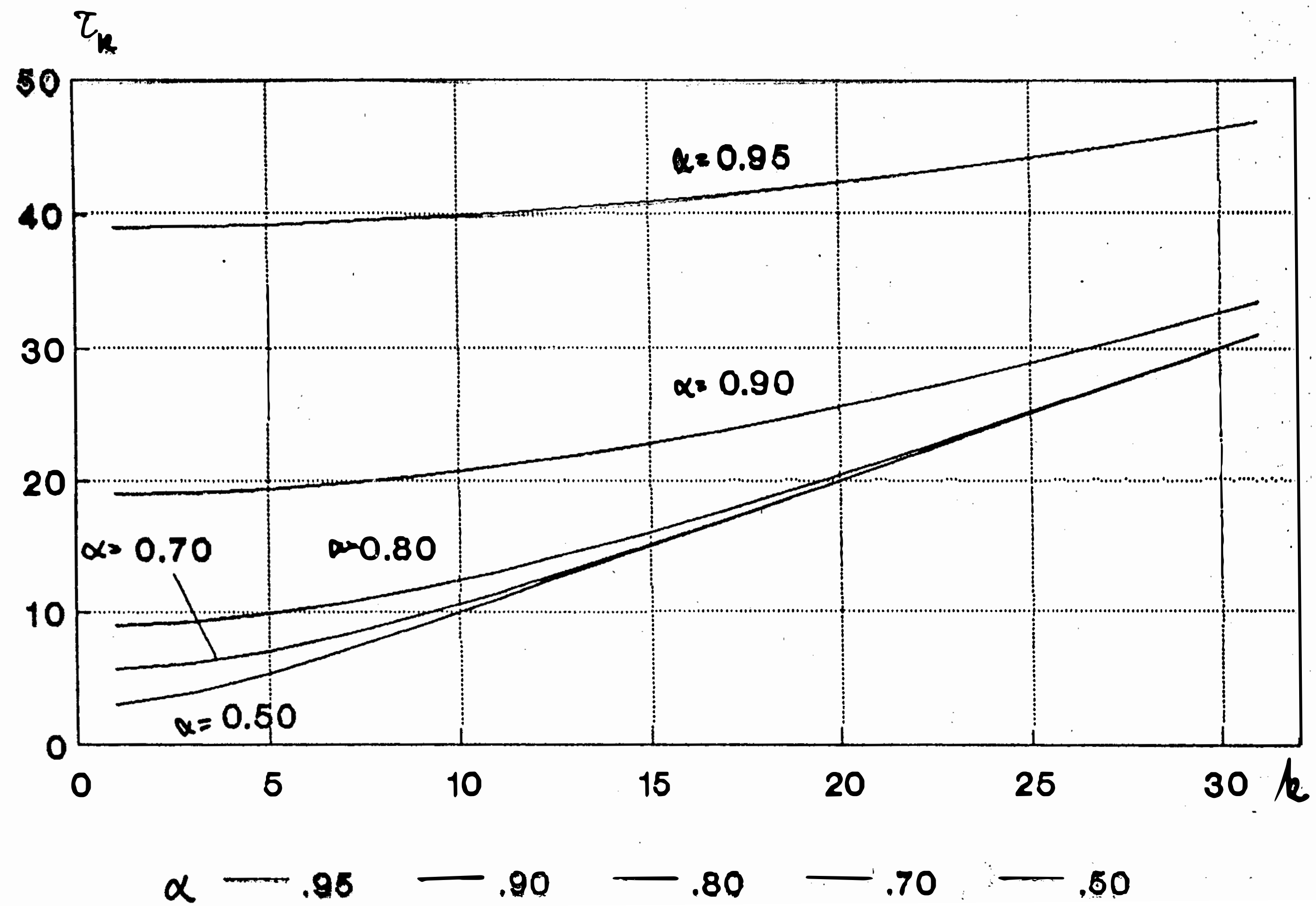
The decorrelation time of $(**)$ is

$$\tau_k = \frac{1+\alpha^k}{1-\alpha^k} \cdot k \geq k$$

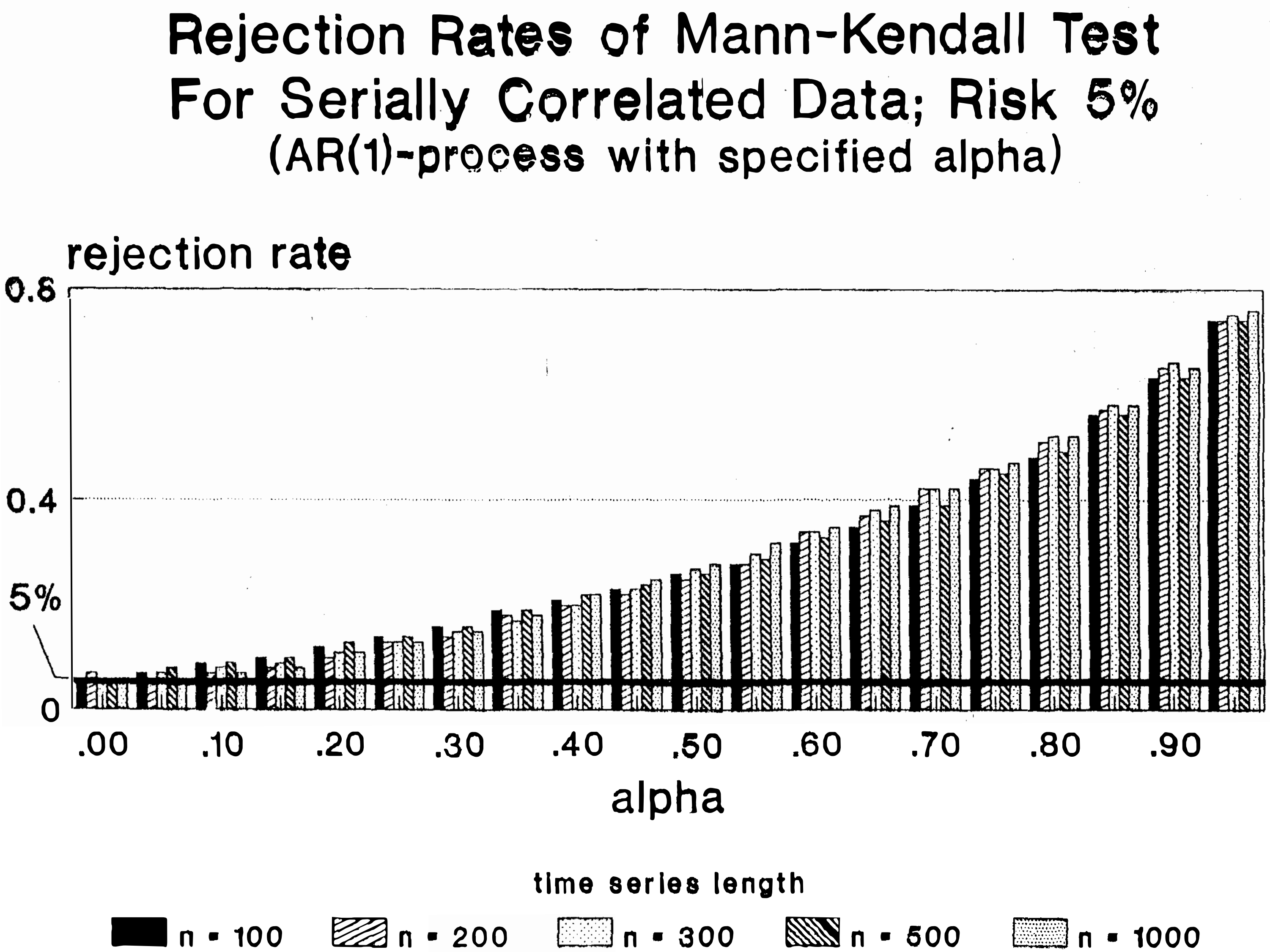
and

$$\lim_{k \to \infty} \frac{\tau_k}{k} = 1$$

Thus, for sufficiently large time increments the decorrelation time is equal to the time increment independent of the "memory" α of the system.

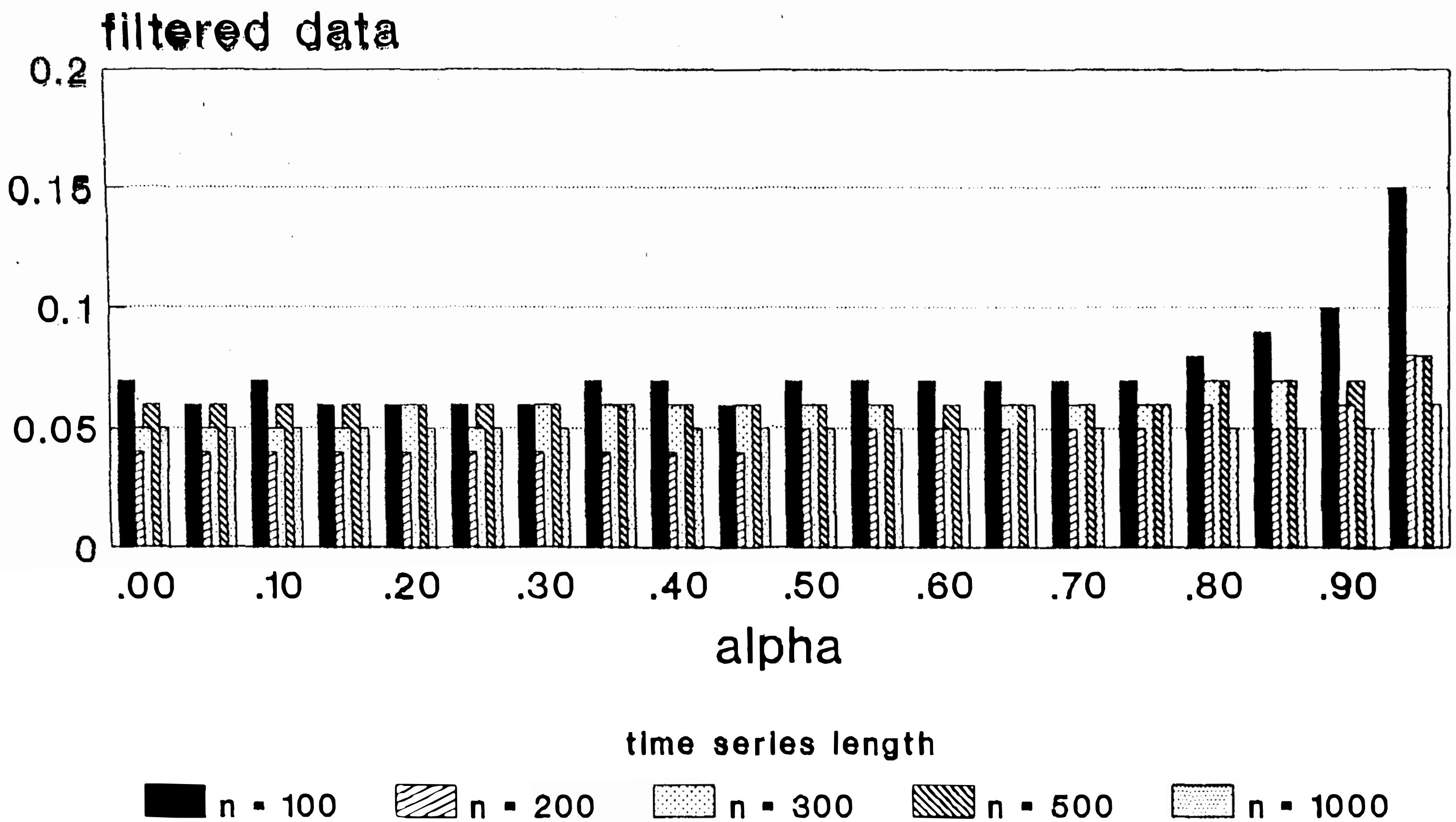


τ_k
50
40
30
20
10
0
0
5
10
15
20
25
30
k
α= 0.95
α= 0.90
α= 0.80
α= 0.70
α= 0.50
α .95 .90 .80 .70 .50



Kulkarni & von Storch, 1995

Rejection Rates of Mann-Kendall Test For Serially Correlated Data; Risk 5%
(AR(1)-process with specified alpha)



Kulkarni & von Storch, 1995

CONCLUSION

Statistics is ...

not **a** *Wunderwaffe* **to extract a wealth of information from a limited sample of observations**

but an indispensable tool in the evaluation of limited empirical evidence

For extracting more information from a data set about the underlying structure assumption about the underlying structure are to be made. In general, such assumptions have to be justified by additional information unrelated to the data (for instance from numerical experimentation or theoretical reasoning).