Cross Correlation

Dr. Zhi Wang
Department of Earth and Environmental Sciences
California State University, Fresno

Descriptive

- Autocorrelation considers only one kind of measurement.
- Cross Correlation (CC) considers relationships between two different kinds of measurements.
- CC reveals over what distances in space or time the two kinds of measurements are related to each other.
- The concepts behind covariance and correlation coefficient (neglecting measurement locations) are used here to determine the relation between two variables observed at different locations.

Relevant questions

- Is it necessary to sample both kinds of measurements at identically the same locations (time)?
- How far apart (in distance and time) can two kinds of measurements be correlated with each other?
- How does the sample volume affect the cross correlation length?
- Over what distance are properties sufficiently cross correlated to allow meaningful spatial interpolation between their measured values, e.g., for mapping.

Cross Correlation Coefficient $r_c(h)$

$$\text{Re}member: \textit{Autocorrelation coefficient}$$

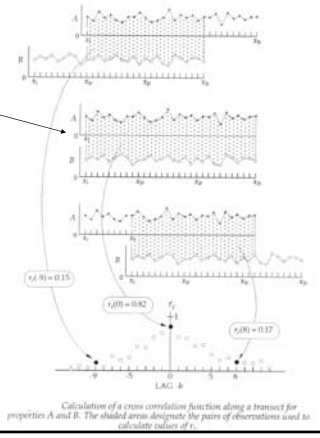
$$r(h) = \frac{\text{cov}\left[A_i(x), A_i(x+h)\right]}{\sqrt{\text{var}\left[A_i(x)\right]}\sqrt{\text{var}\left[A_i(x+h)\right]}}$$

$$\textit{Cross correlation coefficient}\,(r_c)$$

$$r_c(h) = \frac{\text{cov}\left[A_i(x), B_i(x+h)\right]}{\sqrt{\text{var}\left[A_i(x)\right]}\sqrt{\text{var}\left[B_i(x+h)\right]}}$$

Graphic Indications

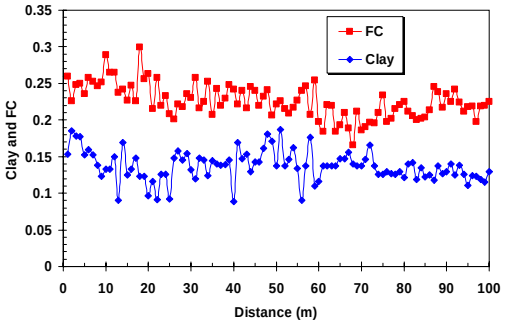
- Leg h=0, $r_c = r$ (the linear correlation coefficient obtained using classical statistical procedures).
- h=1, $r_c(1)$ for the nearest neighbors in the forward direction, for pairs $[A(x_i),B(x_{i+1})]$.
- h=-1, $r_c(-1)$ for the nearest neighbors in the backward direction, for pairs $[A(x_i),B(x_{i-1})]$.
- h=-9 for pairs $[A(x_i),B(x_{i+9})]$



Calculation of a cross correlation function along a transect for properties A and B. The shaded areas designate the pairs of observations used to calculate values of r_c

Clay content and Field Capacity (FC)

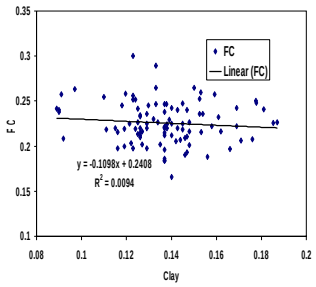
Spatial related? Best leg length h=?



Clay content and Field Capacity (FC)

at h=0 – 20m

leg	CCF	leg	CCF
0	-0.08		
1	0.039	-1	0.034
2	0.057	-2	-0.075
3	0.002	-3	-0.058
4	0.15	-4	-0.032
5	0.226	-5	-0.048
6	0.163	-6	-0.054
7	0.102	-7	0.013
8	0.367	-8	-0.013
9	0.068	-9	-0.019
10	0.124	-10	0.14
11	0.036	-11	-0.082
12	0.091	-12	-0.101
13	-0.048	-13	0.026
14	0.004	-14	0.042
15	0.086	-15	0.035
16	-0.049	-16	0.005
17	0.053	-17	0.043
18	-0.081	-18	-0.05
19	-0.21	-19	0.022
20	-0.1	-20	0.024



Significance of Cross Correlation

- More difficult to judge the significance of Cross correlation compared to Autocorrelation.
- Assumptions are: both A and B are normally distributed, $A(x_i)$ and $B(x_i)$ are independent of each other.
- The t-test can be used to tell if A and B are spatially cross correlated:

$$t = r_c\sqrt{\frac{n^* - 2}{1 - r_c}}$$

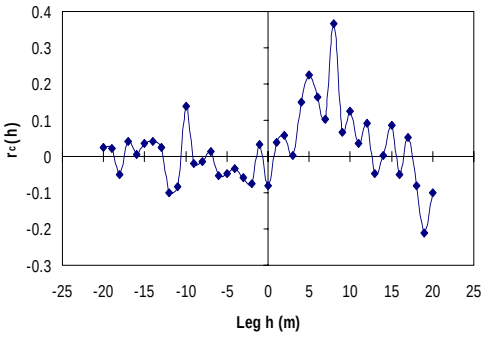
Critical t_c (*EXCEL FUNCTION*):

t_c =TINV(α, n-2) n*= # of overlaped positions used for r_c.

If $t < t_c$, *cross – correlated at significant level of* $1-\alpha$

If $t > t_c$, *not cross – correlated at significant level of* $1-\alpha$

Cross Correlation Function (CCF) of Clay vs FC



Research Topics

- Cross Correlation Functions (CCF) change in time.
- An optimum sampling distance for one time may be inadequate for other times.
- Expect varied CCFs for different physical processes at different times.
- Auto- and CCF ranges reveal the nature and impact of particular processes.

Computer Exercises

- Use MiniTab (or other software) and the provided dataset to find Cross Correlograms of a selected pair of variables.
- Use t-test to determine if the selected two attributes are significantly cross correlated at 1-α = 95% (at what leg length?).