



How to Calculate Covariance using a dataTaker

Covariance is a commonly used calculation used in a wide number of areas from microclimate flux and carbon sequestration, aquaculture, prediction of stock market to measurement of the southern oscillation index. It is a method of measuring how strongly variable are related to each other.

If the result of the covariance calculation is zero then the two variables are not related to one another. When the result is positive then the two variables have moved in the same direction. The larger the result the more strongly related the two variables are. If the result is negative then the two variables have a negative relationship with one another, that is are moving in opposite directions.

Traditionally the formula for covariance is given as:

$$CoVar(A, B) = \frac{1}{n} \sum_{i=1}^n [(A_i - \bar{A}) \cdot (B_i - \bar{B})]$$

Where :

n = Number of samples

\bar{A} = Average of variable A

\bar{B} = Average of variable B

A_i = a single sample of variable A

B_i = a single sample of variable B

Inspection of this formula reveals that we need the average value of each variable before we can calculate the variation of each individual sample is from the average reading. This is computationally intensive and tedious to implement.

It can be shown that the covariance formula can be simplified to:

$$CoVar (A, B) = Ave(A \cdot B) - Ave(A) \cdot Ave(B)$$

On inspection it is soon seen that, as the average of a series of number is sum of the sample divided the number of samples taken, this is far easier to implement from a computational standpoint.

Instead of storing each sample for latter processing we only need to keep running sums of variables A , B and $A \cdot B$ and keep a count of the number of samples taken.

A simple dataTaker program can be implemented to do this. Below is an example which is compatible with the DT80 Series data loggers.



BEGIN "COVAR"

```

=====
' Schedule A (Read Variables)
' - Runs every 1 second
=====

```

RA1S LOGOFFA

```

1V(=1CV,W)           'Read the A and assign to 1CV
2V(=2CV,W)           'Read the B and assign to 2CV
3CV(W)=3CV+1CV       '3CV = Sum of A
4CV(W)=4CV+2CV       '4CV = Sum of B
5CV(W)=5CV+(1CV*2CV) '5CV = Sum A*B
6CV(W)=6CV+1         'Number of samples

```

```

=====
' Schedule B (Reporting Schedule)
' - Reports the Covariance
=====

```

RB("B:",ALARMS:OV:10KB,DATA:OV:1MB)10S LOGONB

```

'Calculate Covar(A,B) and save
7CV("Covariance")=(5CV/6CV)-((3CV/6CV)*(4CV/6CV))
'Reset the variables and start again
1..7CV(W)=0

```

END