Validation and Evaluation

- Effective and transparent

O :
The reference data set, (trustworthy observations or validated model results)

P :
The data set to be evaluated (new model results, ect)

$$\begin{bmatrix} i=1,\ j=1 & \cdots & i=1,\ j=end \\ \vdots & & \vdots \\ i=end,\ j=1 & \cdots & i=end,\ j=end \end{bmatrix}$$

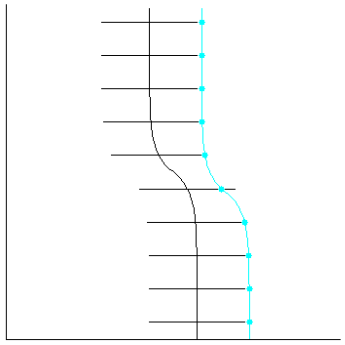
$$\begin{bmatrix} i=1,\ j=1 & \cdots & i=1,\ j=end \\ \vdots & & \vdots \\ i=end,\ j=1 & \cdots & i=end,\ j=end \end{bmatrix}$$

i, j : Index (depth, time, parameter ect.)

Relevant aspects of the data set?

- We want to evaluate climate.
- We want to take into account processes in, and the state of, both surface and basin waters.

Vertical mean profiles
during the evaluation period.



Two objective quality metrics

- Two aspects of data comparison

r :

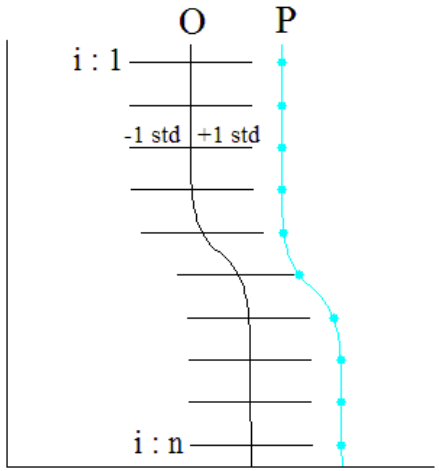
The correlation coefficient, an evaluation of “shape”.

$$r = \frac{\sum_{i=1}^{n}\left(P_i - \overline{P}\right)\left(O_i - \overline{O}\right)}{\sqrt{\sum_{i=1}^{n}\left(P_i - \overline{P}\right)^2 \sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2}}$$

C :

The mean of the cost function, an evaluation of “level”.

$$C = \frac{\sum_{i=1}^{n}\left|\frac{P_i - O_i}{std\left(O_i\right)}\right|}{n}$$



O : The reference data set, (trustworthy observations or validated model results)

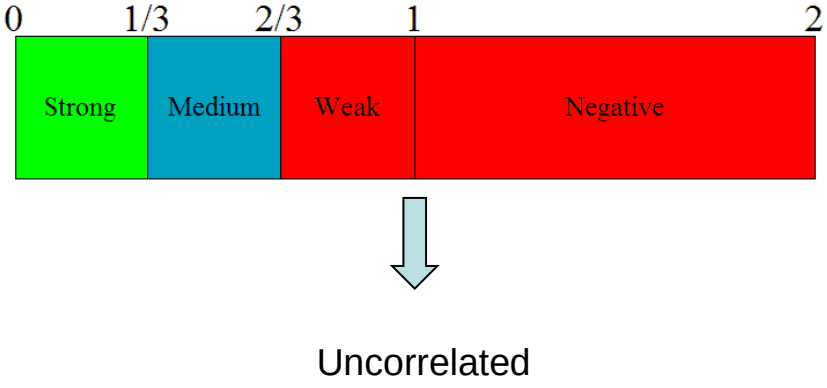
P : The data set to be evaluated (new model results, ect)

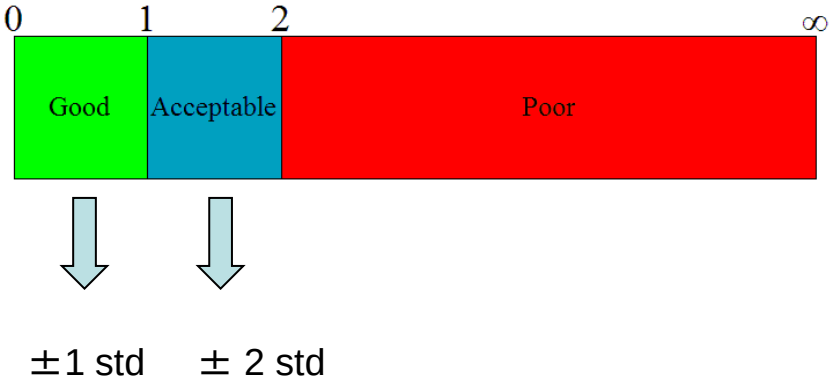
i : Index over which the evaluation is preformed (depth, time, ect.)

Defining what's good

(1-*r*) :
The correlation coefficient ,
expressed in the interval 0-2.

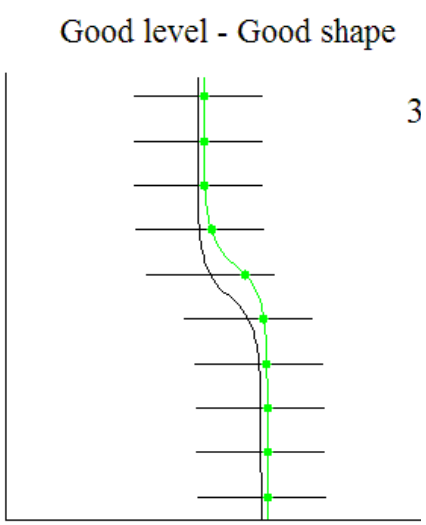
C :
The mean of the cost function.



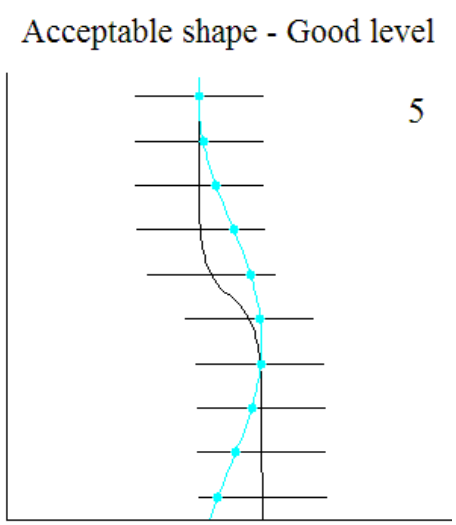


Examples of profile shape

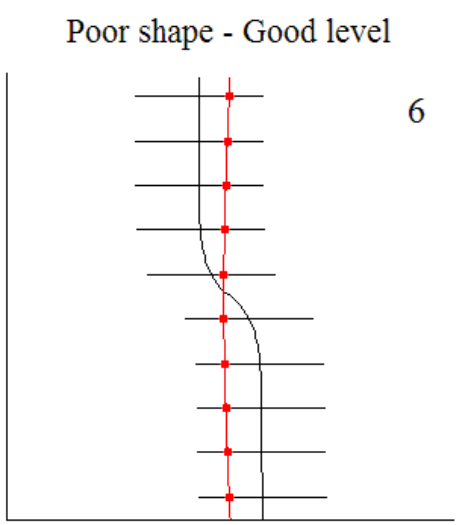
r ~ 0.05



r ~ 0.42

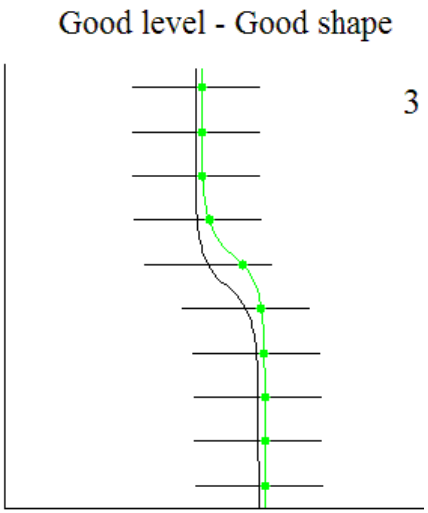


r ~ 0.98

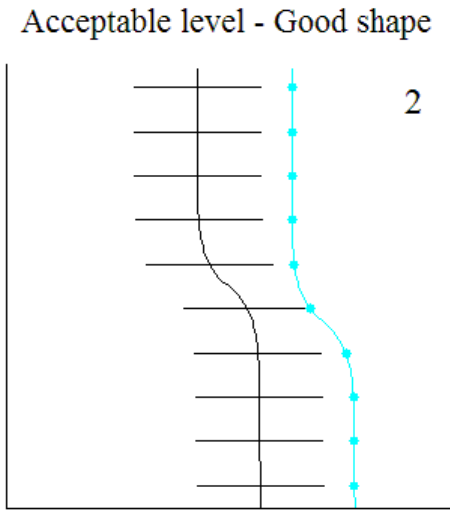


Examples of profile level

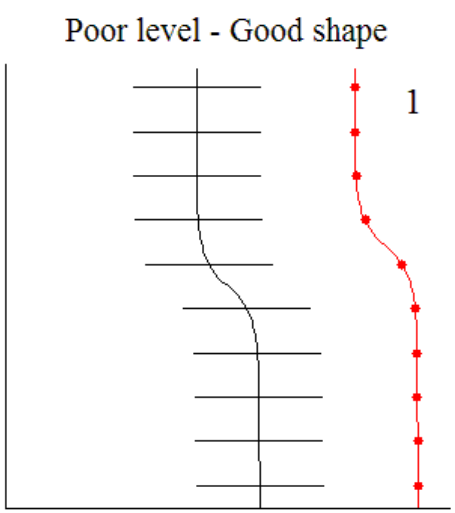
C ~ 0.17



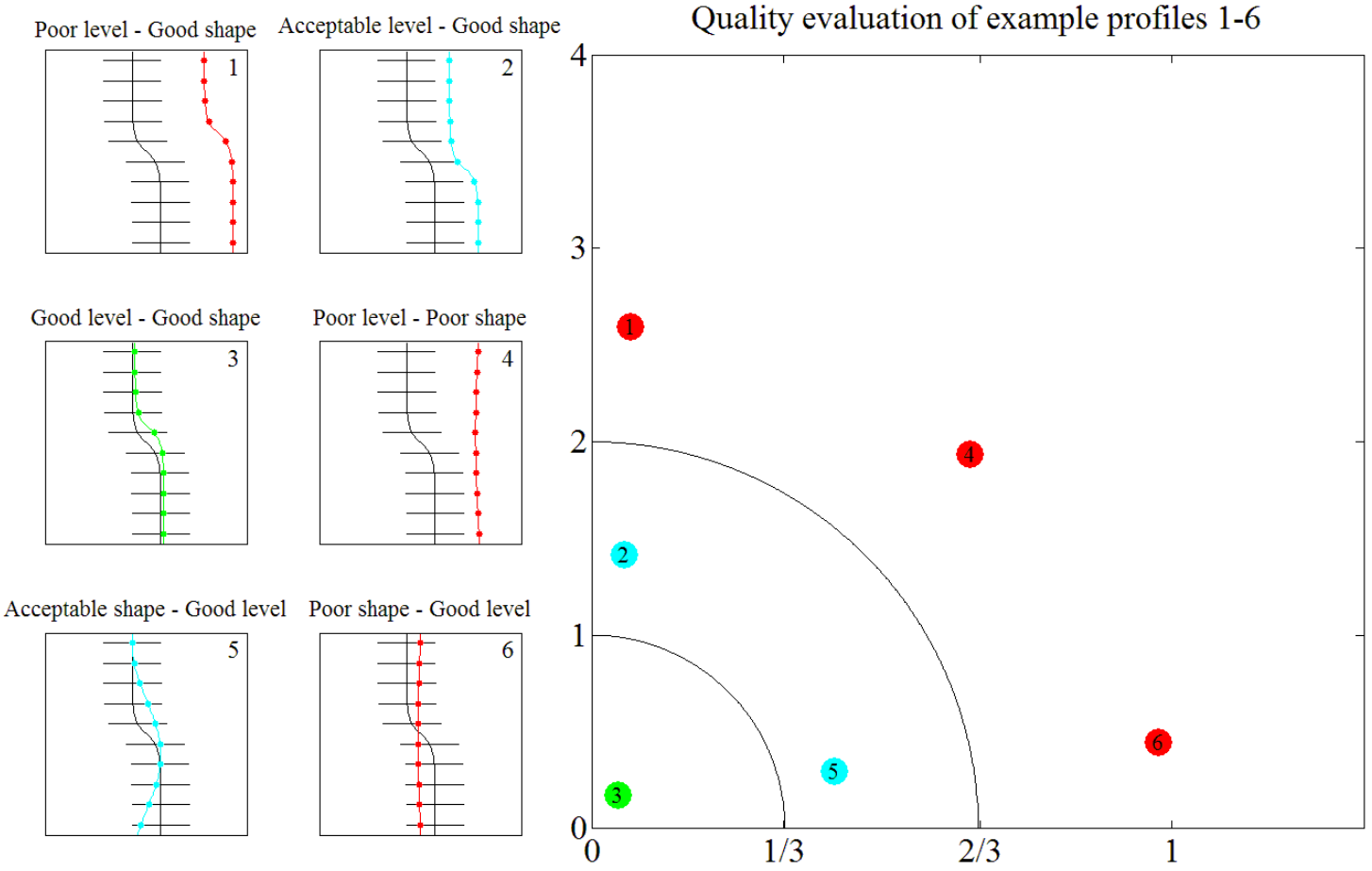
C ~ 1.4



C ~ 2.6



A simple graphic representation



Summation to a single dot

7 key parameters are evaluated (volume weighted climatic profiles): S,T, NO3, PO4, O2, AT and pH

Since *r* and *C* are dimensionless, a mean can be taken across all the parameters.

This reduces evaluation of the entire data set to two numbers.
... Or one single dot.

Validation of the scenario model (PROBE-Baltic 3.0)