The MobiView graphical user interface

Magnus Dahler Norling

November 15, 2019

Contents

1 Introduction

MobiView is a graphical user interface that can load any model built using the Mobius dll interface. This document is a work in progress.

2 Goodness of fit statistics

Most of the goodness-of-fit statistics are implemented following [1]. Further properties of the various measures can be found in that paper.

Let $o = \{o_i\}_{i\in I}$ be the observed timeseries, and let $m = \{m_i\}_{i\in I}$ be the modelled timeseries. The set I of comparison points is the set of all timesteps in the GOF interval where both series have a valid value. For instance, the observed timeseries can have missing values, so the timesteps corresponding to the missing values will not be considered when evaluating goodness-of-fit. The GOF interval is the entire model run interval unless something else is specified by the user. Let

$$\overline{m} = mean(m)$$

denote the mean of a timeseries.

2.1 Common data points

The common data points is the size of the set of comparison points I, denoted $|I|$.

2.2 Mean error (bias)

The mean error is

$$\overline{o-m} = \bar{o} - \overline{m} = \frac{1}{|I|}\sum_{i\in I}(o_i - m_i)$$

For fluxes or flows, the mean error is related to the discrepancy in mass balance.

2.3 MAE

MAE is the mean absolute error

$$\frac{1}{|I|}\sum_{i\in I}|o_i - m_i|,$$

where $|\cdot|$ denotes the absolute value of a number.

2.4 RMSE

RMSE is the root mean square error

$$\sqrt{\frac{1}{|I|}\sum_{i\in I}(o_i - m_i)^2}$$

2.5 N-S

N-S is the Nash-Sutcliffe efficiency coefficient

$$1 - \frac{\sum_{i\in I}(o_i - m_i)^2}{\sum_{i\in I}(o_i - \bar{o})^2}$$

This coefficient takes values in $(-\infty, 1]$, where a value of 1 means a perfect fit, while a value of 0 means that the modeled series is a no better fit than the mean of the observed series.

2.6 log N-S

log N-S is the same as N-S, but where o_i is replaced by $\ln(o_i)$ and m_i replaced by $\ln(m_i)$ for each $i \in I$, where ln denotes the natural logarithm.

$$1 - \frac{\sum_{i\in I}(\ln(o_i) - \ln(m_i))^2}{\sum_{i\in I}(\ln(o_i) - \overline{\ln(o)})^2}$$

This coefficient behaves similarly to N-S, but is less sensitive to errors during high values.

2.7 r2

r^2 is the coefficient of determination

$$\left(\frac{\sum_{i\in I}(o_i-\overline{o})(m_i-\overline{m})}{\sqrt{\sum_{i\in I}(o_i-\overline{o})^2}\sqrt{\sum_{i\in I}(m_i-\overline{m})^2}}\right)^2$$

This coefficient takes values in $[0,1]$.

2.8 Idx. of agr.

The index of agreement is

$$1-\frac{\sum_{i\in I}(o_i-m_i)^2}{\sum_{i\in I}(|m_i-\overline{o}|+|o_i-\overline{o}|)^2}$$

2.9 Spearman's RCC

Spearman's rank correlation coefficient [2] is computed as follows: For a timeseries $x = x_{i i\in_I}$, let $\mathrm{rank}(x_i)$ be the index of x_i (starting from 1) in the list $\mathrm{sort}(x)$, where $\mathrm{sort}(x)$ is x sorted from smallest to largest. The rank correlation coefficient can then be computed as

$$1-\frac{6\sum_{i\in I}(\mathrm{rank}(o_i)-\mathrm{rank}(m_i))^2}{|I|(|I|^2-1)}$$

The coefficient takes values in $[-1,1]$. If the value is 1, the modeled series is a (positively) monotone function of the observed series.

References

[1] P. Krause, D. P. Boyle, and F. Bäse. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5:89–97, 2005.

[2] C. Spearman. The proof and measurement of association between two things. *American Journal of Psychology*, 15:72–101, 1904.