

# Goodness-of-fit Measures to Compare Observed and Simulated Values with hydroGOF

Mauricio Zambrano-Bigiarini

Aug 2011

## 1 Installation

Installing hydroGOF:

```
> install.packages("hydroGOF")
```

## 2 Setting Up the Environment

1. Loading the *hydroGOF* library, which contains data and functions used in this analysis.

```
> library(hydroGOF)
```

2. Loading observed streamflows of the Ega River (Spain), with daily data from 1961-Jan-01 up to 1970-Dec-31

```
> require(zoo)
> data(EgaEnEstellaQts)
> obs <- EgaEnEstellaQts
```

3. Generating a simulated daily time series, initially equal to the observed values (simulated values are usually read from the output files of the hydrological model)

```
> sim <- obs
```

4. Computing the numeric goodness-of-fit measures for the "best" (unattainable) case

```
> gof(sim=sim, obs=obs)
```

```
      [,1]
ME      0
MAE     0
MSE     0
RMSE    0
NRMSE % 0
PBIAS %  0
```

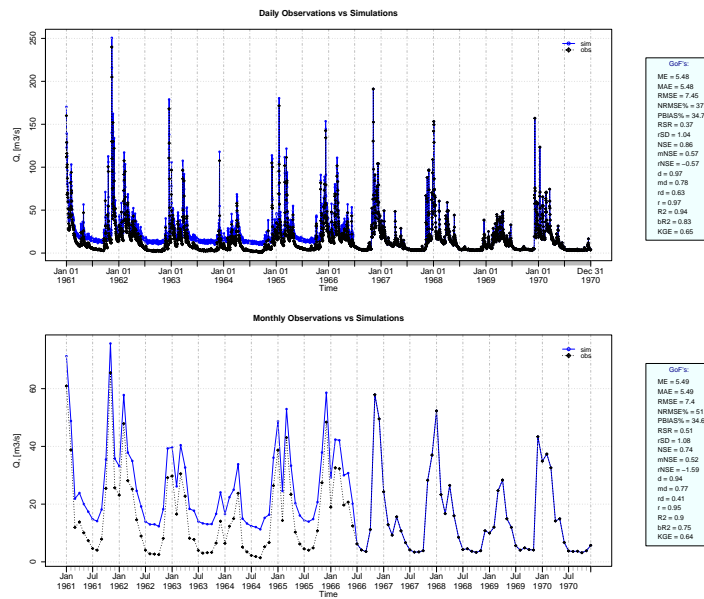
RSR	0
rSD	1
NSE	1
mNSE	1
rNSE	1
d	1
md	1
rd	1
cp	1
r	1
R2	1
bR2	1
KGE	1

- Randomly changing the first 2000 elements of 'sim', by using a normal distribution with mean 10 and standard deviation equal to 1 (default of 'rnorm').

```
> sim[1:2000] <- obs[1:2000] + rnorm(2000, mean=10)
```

- Plotting the graphical comparison of 'obs' against 'sim', along with the numeric goodness-of-fit measures for the daily and monthly time series

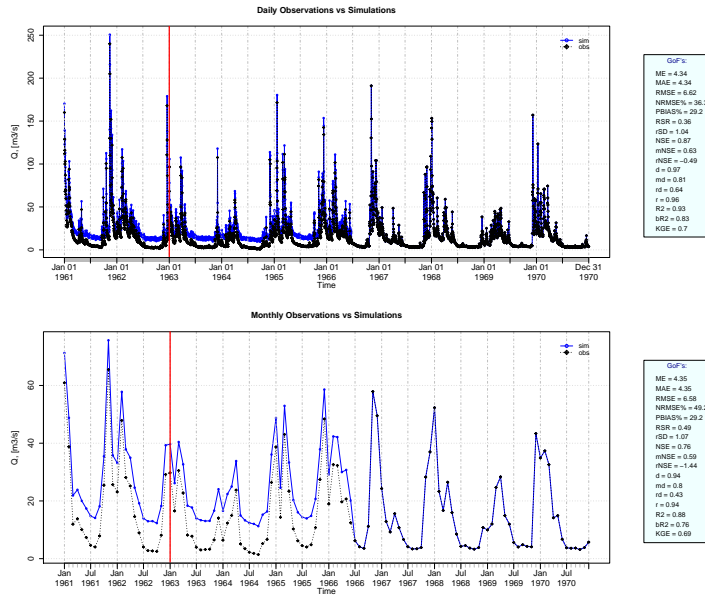
```
> ggof(sim=sim, obs=obs, ftype="dm", FUN=mean)
```



### 3 Removing Warm-up Period

- Using the first two years (1961-1962) as warm-up period, and removing the corresponding observed and simulated values from the computation of the goodness-of-fit measures:

```
> ggof(sim=sim, obs=obs, ftype="dm", FUN=mean, cal.ini="1963-01-01")
```



2. Verification of the goodness-of-fit measures for the daily values after removing the warm-up period:

```
> sim <- window(sim, start=as.Date("1963-01-01"))
> obs <- window(obs, start=as.Date("1963-01-01"))
> gof(sim, obs)
```

```
      [,1]
ME      4.34
MAE     4.34
MSE    43.83
RMSE     6.62
NRMSE % 36.30
PBIAS % 29.20
RSR      0.36
rSD      1.04
NSE      0.87
mNSE     0.63
rNSE    -0.49
d        0.97
md       0.81
rd       0.64
cp       0.44
r        0.96
R2       0.93
bR2      0.83
KGE      0.70
```

## 4 Analysis of the Residuals

1. Computing the daily residuals (even if this is a dummy example, it is enough for illustrating the capability)

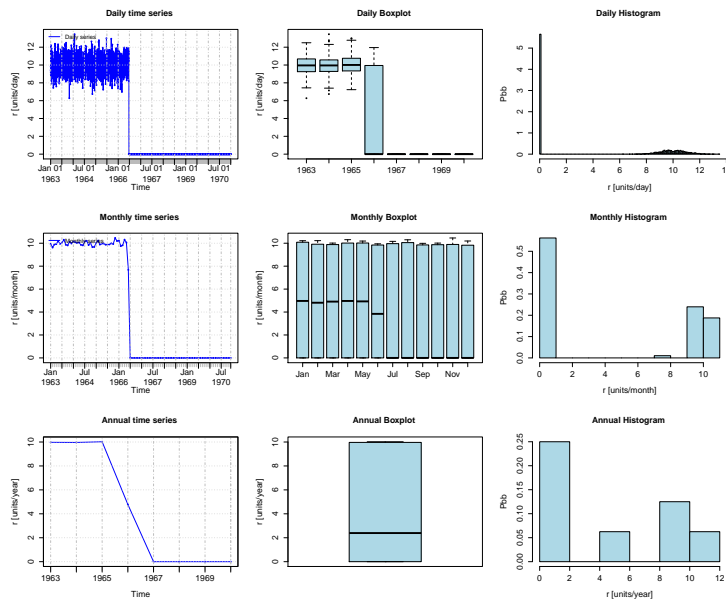
```
> r <- sim-obs
```

2. Summarizing and plotting the residuals (it requires the hydroTSM package):

```
> library(hydroTSM)
> smry(r)
```

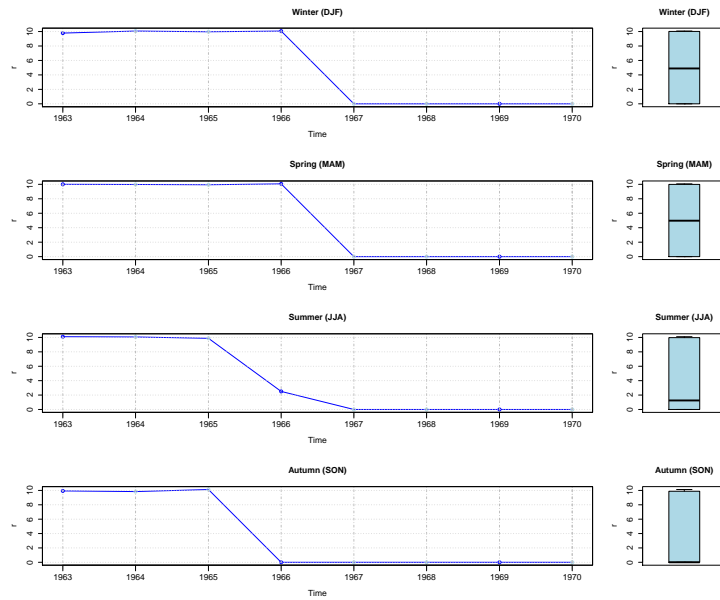
	Index	r
Min.	1963-01-01	0.0000
1st Qu.	1964-12-31	0.0000
Median	1966-12-31	0.0000
Mean	1966-12-31	4.3430
3rd Qu.	1968-12-30	9.7650
Max.	1970-12-31	13.4600
IQR	<NA>	9.7649
sd	<NA>	4.9978
cv	<NA>	1.1509
Skewness	<NA>	0.3160
Kurtosis	<NA>	-1.8324
NA's	<NA>	2.0000
n	<NA>	2922.0000

```
> # daily, monthly and annual plots, boxplots and histograms
> hydroplot(r, FUN=mean)
```



### 3. Seasonal plots and boxplots

```
> # daily, monthly and annual plots, boxplots and histograms  
> hydroplot(r, FUN=mean, pfreq="seasonal")
```



This tutorial was built under:

- [1] "x86\_64-redhat-linux-gnu (64-bit)"
- [1] "R version 2.15.0 (2012-03-30)"
- [1] "hydroGOF 0.3-3"