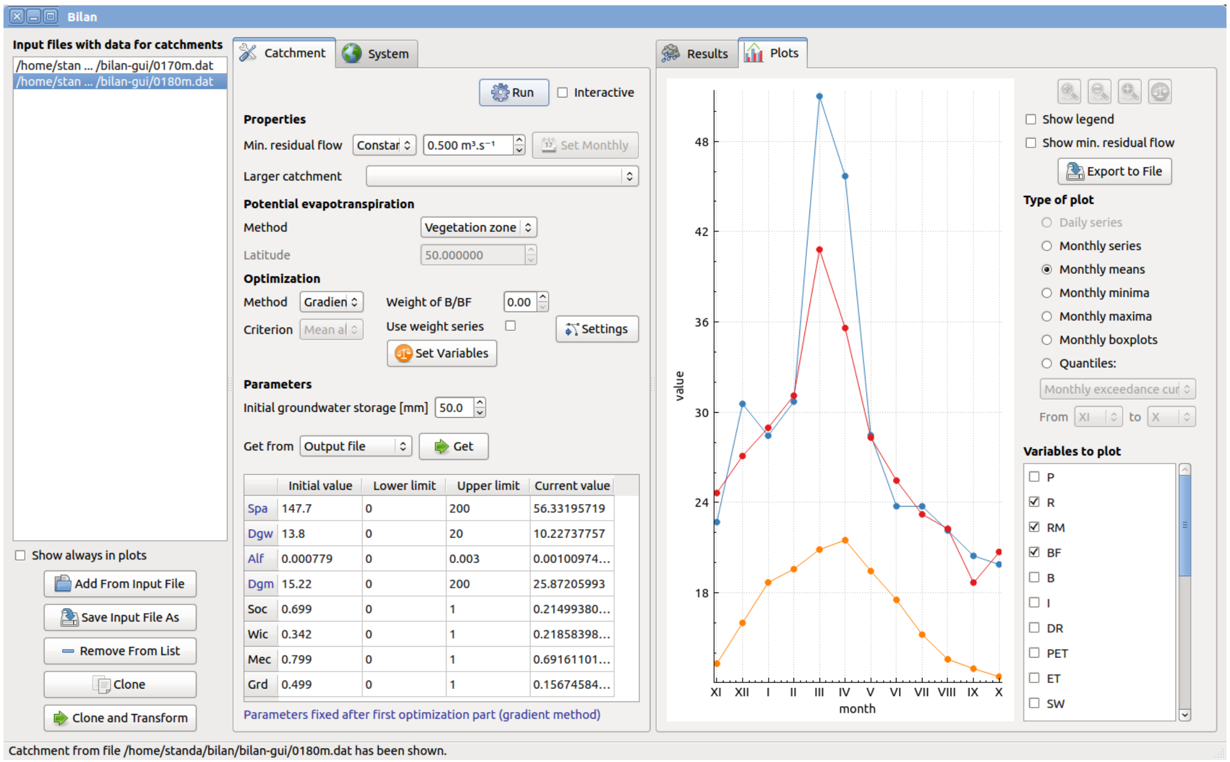
Bilan water balance model

Manual



version 2015-06-23

T. G. Masaryk Water Research Institute in Prague

August 18, 2015

Contents

Introduction

The Bilan model has been developed to simulate components of the water balance in a catchment. The structure of the model is constituted by a system of relationships describing basic principles of water balance on the ground, in the aeration zone of the soil including the effect of vegetation cover, and in the groundwater zone. Air temperature is used as an indicator of energy conditions, which significantly affect the water balance components. The time resolution of the model is one day or one month.

The input data used for water balance computation include the daily or monthly series of basin precipitation, air temperature and relative air humidity (optional). To calibrate the parameters of the model (applying the optimisation algorithm), simulated and observed daily or monthly runoff series at the outlet from the basin are used.

The model simulates the time series of daily or monthly potential evapotranspiration, actual evapotranspiration, soil infiltration and recharge from the soil to the aquifer. The amounts of water stored in the snow pack, in the soil and in the aquifer are also simulated for each time step. All these hydrological variables apply to the whole catchment. The total runoff consists of two or three components, i.e. direct runoff, interflow (for the monthly time step only) and base flow.

The model makes use of six (daily time step) or eight (monthly time step) free parameters and of an optimisation algorithm for their calibration in the gauged basins. The optimisation is aimed at attaining the best fit between the observed and simulated runoff series, for which several optimisation criteria are available.

The temperature and eventually the relative humidity are used to calculate the potential evapotranspiration. Temperature is also used to distinguish between winter and summer conditions (the regime type). If there is a snow pack on the ground, a snow storage algorithm and a melting algorithm are applied. Melting snow and rainfall infiltrate into the soil. The infiltrated water stored in the soil can later be extracted by agricultural crops or natural vegetation. The crops or vegetation utilise the soil moisture at a certain potential rate (potential evapotranspiration) as long as there is sufficient water in the soil. If there is not enough water in the soil the actual evapotranspiration will drop below the potential rate. During wet periods, when the precipitation exceeds the potential evapotranspiration, the surplus is used to replenish the soil water storage. Percolation from the soil occurs when the amount of water stored in the soil exceeds the maximum soil storage capacity. Water percolating from the soil can follow a quick route towards a stream due to interflow (monthly type only) or as recharge, or it may follow a slow route through the aquifer. A third (second) streamflow component, i.e. surface runoff, may also occur in case of a high rainfall.

User's manual

1 Installation

The core of the BILAN model is written in the C++ programming language. Two interfaces to this core are available, including a graphical user interface (GUI) based on Qt5 libraries and a package for the R statistical and programming environment.

In this manual, the core features are described in sections of text which are not highlighted.

1.1 Graphical user interface

For Windows, the user is provided with the bilan_VERSION.exe file, where the version number is based on date of release. For its execution, 11 .dll files are needed, including 10 files located in the same directory (e.g. libgcc_s_dw2-1.dll, mingwm10.dll, QtCore5.dll or QtGui5.dll) and the qwindows.dll located in the platforms subdirectory.

For Linux, an executable binary is also available.

In this manual, the GUI-specific sections are highlighted in yellow.

1.2 R package

The R 2.12 or higher (2.15 or higher is recommended) with the Rcpp package installed (version 0.9.10 or higher) is required for the Bilan package.

The package bilan_VERSION.tar.gz (for Linux) or bilan_VERSION.zip (for Windows) can be installed into the R environment in a standard way routinely used to install local packages (e.g. using the `install.packages` command). In this manual, only function names are referred (highlighted in blue); the detailed description is contained in R-help associated with each individual function.

2 User interface

The application user interface consists of three main sections:

- the list of files with catchment data and tools for their management,
- catchment properties and model settings for the catchment currently selected or for a system of catchments represented by the *Catchment* and *System* tabs,
- results based on the model settings represented by the *Results* and *Plots* tabs.

The user interface sections can be resized or hidden completely by using vertical handles that divide them.

To quit the application, choose *Bilan* → *Quit* in the main menu or use the keyboard shortcut Ctrl+Q.

2.1 User interface settings

General values and options chosen in the user interface are automatically saved and restored as it is usual in software applications. This applies to the options on the *Results* and *Plots* tabs as to the general dialogs (*Preferences*, *Load Input/Output File* and *Set Variables*). The location of the saved configuration depends on the type of operating system used: under Windows, it is stored in the registry, under Linux it is in the file $HOME/.config/wri/bilan.conf.

Settings related to a particular catchment selected from the list of catchments (i.e. mostly the optimization settings) can be saved and reused by using the *Tools* → *Save Catchment Settings* and *Tools* → *Load Catchment Settings* menu items or by keyboard shortcuts Ctrl+S and Ctrl+L. The settings are stored in the INI file in a user-defined location.

2.2 Application preferences

Preferences used for whole application are set in the dialog invoked by the *Tools* → *Preferences* menu item or by keyboard shortcut Ctrl+P.

Language change is applied immediately after confirming the dialog. Currently English and Czech localizations are available.

The measurement units that can be set include *Height* (the height of the water column in millimeters, default option) or *Volume* (the flow in $m^3.s^{-1}$ or the total volume per time step in m^3). This option affects all values of the relevant variables including their output to files and visualization in plots. If the *Volume* is chosen, it is not possible to add catchment unless its area is set.

The *Use system file dialogs* ensures that the native file dialogs for given desktop environment are used instead of dialogs provided by the Qt library. Although the native dialogs are usually more suitable because of their integration with the desktop, they can produce displays in a different language (when system localization differs) and may cause problems in the KDE environment when switching the languages.

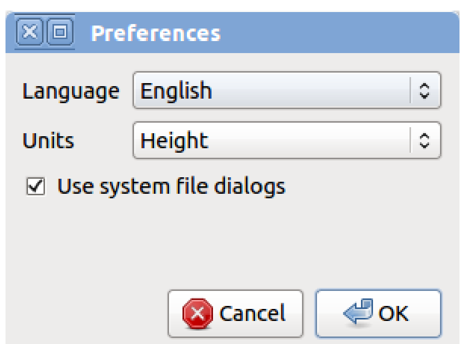


Figure 1: The preferences of Bilan application

3 List of files

The list of files with catchment data is situated at the left-hand side section of the user interface. Below the list, there are tools performing operations on the files.

One of the files can be selected. For the selected catchment, data and plots are shown on the other tabs. The first catchment is automatically added to the system of catchments (see section 7). To show a catchment permanently in the plots, check the *Show always in plots* box.

Buttons below the list allows manipulations with the catchment files.

A new catchment can be added to the list by the *Add From File* button. It is possible to load both input and output files. After selecting one or more files, the dialog *Load Input/Output File* appears for each file to be added. The input file format and the dialog options are described in section 3.1, the output file format is described in section 8.

The input file of a catchment from the list can be saved by pressing the *Save Input File As* button. If no file extension is entered by user in the file dialog, the extension .dat is applied automatically. Note that no information about units is stored in the input file, therefore the units have to be set manually in the *Load Input/Output File* dialog if needed.

A catchment is removed from the list by pressing the *Remove From List* button. The original file with input data remains untouched.

By pressing the *Clone* button, a new catchment is added to the list. The catchment is identical with the catchment being selected when cloning, except that its file name is reset to unsaved_clone. No catchment file is created during cloning. If the catchment is not saved before exiting the application, its data are lost.

```
bil.clone
```

The *Clone and Transform* button invokes the *Transform Input Data* dialog where the transforming coefficients can be specified for each input variable (see details in section 3.2). After accepting the dialog, a new catchment is added to the list with its input time series transformed and its name reset to unsaved_transformed_clone. Regarding saving of the new catchment, the same rules as for clone function applies.

3.1 Input or output data loading

Data of the time-series of variables can be loaded to the Bilan model from two types of plain text files:

- input files containing only input variables without their names,
- output files containing all variables including their names and parameter values.

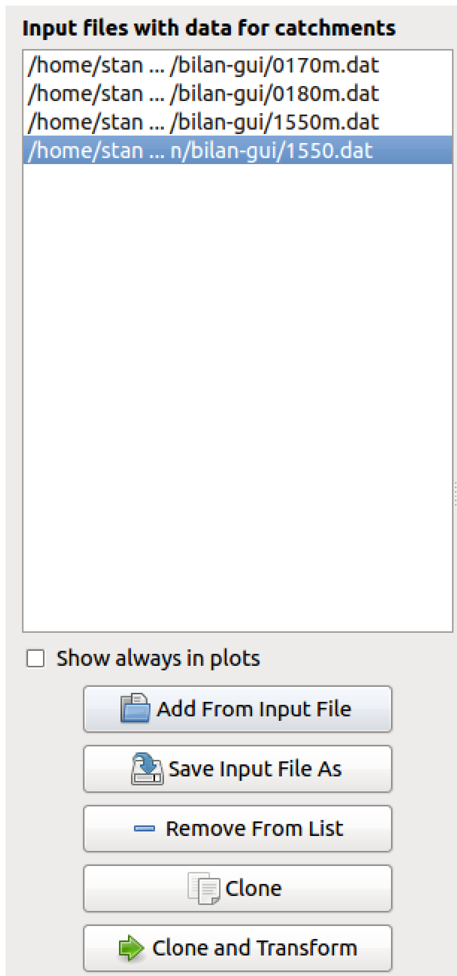


Figure 2: List of files with data for catchments

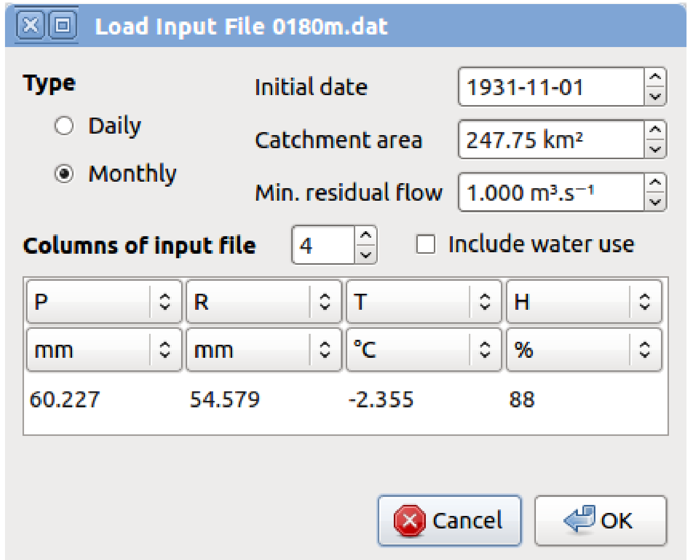


Figure 3: The *Load Input File* dialog

Both types have to follow a given file format. The input file format is described further in this section. The output files are covered by section 8, however these files are produced when saving the model and it is not recommended to edit them manually.

The filename (including the file path) must not contain any special characters or diacritical marks, otherwise the file cannot be loaded.

> When the *Add From File* button is pressed, file dialog to select one or more input or output files appears. After that, the *Load Input/Output File* dialog for each selected file appears, allowing further specification and description of catchment data from the file.
>
> The model type – daily or monthly – can be chosen under *Type*. The type is correctly set for output files automatically whereas an initial guess of type based on the first time step of data is made for input files.

> `bil.new`, `bil.read.file`

Catchment metadata

The initial date of the time series is defined at the first line of the input file formatted as YYYY MM DD (separated by whitespaces). Where the day is not specified it is assumed to be the first day of the month. Where no month is specified it is assumed that the year specified refers to the water year beginning on the 1st of November and that the initial time step is the first day or month of the water year; thus the calendar year of the first time step is one less than the specified year.

The initial date at the first line is optionally followed by catchment area in km^2, separated by a whitespace. The value at the first line is interpreted as catchment area if this is the fourth value or if the value is a decimal number separated by a period.

> The initial date, the catchment area and also the threshold value can be specified in the corresponding boxes of the *Load Input/Output File* dialog. These values replace the data loaded from the file. A more advanced setting of the threshold value is available on the *Catchment* tab (see section 4).

> `bil.set.values`, `bil.set.area`

Time series of variables

The second and the following rows of the input file contain daily or monthly data of the input series. The matrix structure of the data uses rows for data from particular time steps of the daily or monthly time series, while columns are used for individual variables. Missing values can be provided as "NA" strings.

The default input variables are set as follows:

1. precipitation P [mm],
2. runoff (observed) R [mm],
3. air temperature T [°C],
4. relative air humidity H [%] (optional in case of the Oudin method for PET estimation or PET as input data),
5. optional series B – this column can contain any time series. It is a component of the output files and can be used in the graphical outputs of the model. If a non-zero weight for baseflow calibration is set, this column is assumed to contain the baseflow series (estimated based on groundwater measurements) used for comparison with the base flow calculated by the model.
6. potential evapotranspiration PET [mm] (optional),

7. for simulations regarding water use, the groundwater withdrawal series POD [mm], the registered surface water withdrawals POV [mm], the unregistered surface water withdrawals PVN [mm] and the water release values VYP [mm].

The table at the bottom of the *Load Input/Output File* dialog shows variables from the file. The number of variables is the same as the number of columns in the file, although it can be reduced in the *Columns of input/output file* box. Three rows are displayed in the table: drop-down lists to choose a variable, drop-down lists to choose a variable measurement unit for the variable and values from the first line of data.

In case of the input file, variables are selected from the list of available input variables. If water use should be considered, the *Include water use* box needs to be checked which extends the list.

The measurement units relate to the values in the input file. If no catchment area is specified and the units in the drop-downs or in application preferences are set to volumes, the catchment data cannot be loaded.

For the output file, variable names are read from the file and cannot be changed. Water use is also set automatically and the corresponding box is disabled. The same applies for measurement units, they are set according to information stored in the file.

On accepting the dialog, the file is read in and is added to the list of files.

`bil.read.file`, `bil.set.values`

Possible error messages associated with opening and reading of the file are as follows:

- ▷ File does not exist: The input file 'FILE NAME' does not exist. This message also appears when the filename contains forbidden characters.
- ▷ A line does not contain all variables (as defined by the number of variables): Incomplete line 'LINE CONTENT' found.
- ▷ Initial date is not defined correctly: Invalid date format.
- ▷ There is not enough columns for the given variables: Number of columns in file 'FILE NAME' is less than number of input variables.

When the number of columns in the file is greater than the number of variables, the remaining columns are ignored and the warning is shown: The input file 'FILE NAME' contains more columns than input variables, some columns will be omitted.

When a variable is chosen for more than one column, only the last one is used and the warning is shown: File 'FILE NAME': Variable VARIABLE is set for more columns, only the last one will be used.

> When it appears that daily data have been loaded for the monthly type (or the other way round), the warning It seems that daily data are loaded for monthly type. (or It seems that monthly data are loaded for daily type.) is displayed.
>
> If no catchment area is inputted but a conversion of measurement units is required, an error message is shown and no catchment data are loaded.

Examples of input file data

Daily time series beginning 1st of November 1961:

```
1961 11 1 131.88
0.08    0.7833  5.6     85
0       0.7833  6.9     91
0.2     0.7833  6.1     78
...     ...     ...     ...
```

Monthly time series beginning 1st of November 1931 (beginning of the water year 1932):

```
1932 247.75
5.957   17.658  1.454   84
63.117  20.115  -3.574  87
60.227  54.579  -2.355  88
...     ...     ...     ...
```

Note: Old-style formatted input files are also supported for backward compatibility, but their use is not recommended. The header of these files indicates the total number of time steps (first row), the number of columns (second row) and the initial water year (third row). When loading data from old-style formatted files, a warning message is shown.

Old-style formatted output files are not supported and cannot be loaded.

Loading of the output file

If the model type does not match the type in the output file, the type is modified automatically.

After loading from the output file, the simple model run is set, i.e. the number of gradient optimization iterations is set to zero.

It is worth noting that the output file does not contain the complete information needed for the model run (e.g. PET method is missing). Therefore, rerunning the model created from the output file may result in different values of variables than those stored in the file.

Validation of data

After the data are read in, the mean annual values (described in section 8) of all available series are calculated, and these are subsequently displayed on the *Results* tab together with the file name, the number of time steps and both the initial and the end dates. This allows the user to check whether the data were correctly read in.

3.2 Input data transformation

When the *Clone and Transform* button is pressed, the *Input Data Transformation* dialog appears. This dialog allows to set values of linear transformation of the input data series, i.e. the constant (an addend) to be added to all members of the selected series and the constant (a factor) by which all members of the series are to be multiplied.

Moreover, the factors and addends can differ within the year. The *Set Monthly for Current Item* button invokes the *Set Monthly Values* dialog, related to the currently selected table cell. The factor or addend for each calendar month can be specified in this dialog. In addition, the monthly values can be saved to the text file and loaded later by using the *Save* and *Load* buttons. Each row in the file contains a monthly value, starting from January.

When such values are set, the text “monthly” is shown in the table cell instead of the value; the monthly coefficients can be further modified (by re-launching the monthly dialog) or changed to a constant value that does not change during the year.

If the interactive mode (see section 5.5) is active, a new catchment is added to the list when the input transformation dialog is launched. Then, any change to the coefficients (including changes in the monthly dialog) triggers calculations for the new catchment, with the input data transformed. If the transformation dialog is accepted, the new catchment remains on the list, otherwise it is deleted.

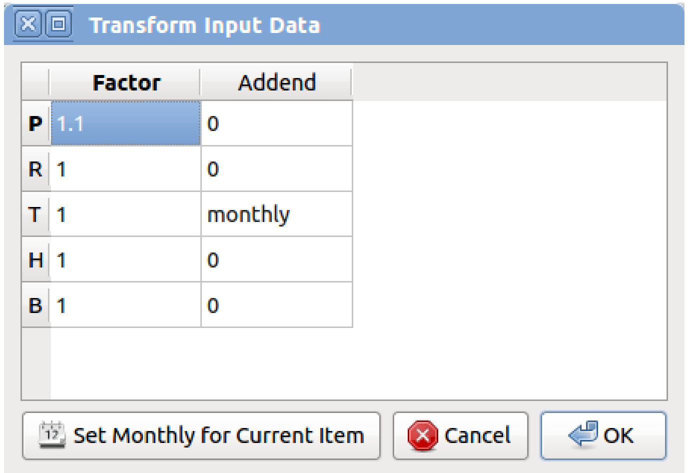


Figure 4: The *Transform Input Data* dialog

There are other opportunities for transforming or creating input data in the *Set Variables* dialog (see section 4.1).

4 Catchment properties

The *Catchment* tab provides access to catchment properties and the settings needed for running the model. The properties and settings (e.g. optimization type, parameter limits) shown refer to the catchment currently selected from the list of files and may differ for different catchments. This section describes the catchment properties; other settings are described in section 5.

The threshold value (representing e.g. the minimum residual flow) can be shown in plots (see section 9) and is used to calculate deficit volumes when water use variables have been loaded. There are three options for this setting:

- *Not set*,
- *Constant* which is also set when a threshold value is contained in the input file for the catchment or has been entered in the *Load Input/Output File* dialog,
- *Monthly* which allows to invoke the dialog for entering monthly values by the *Set Monthly* button (the dialog is described in section 3.2).

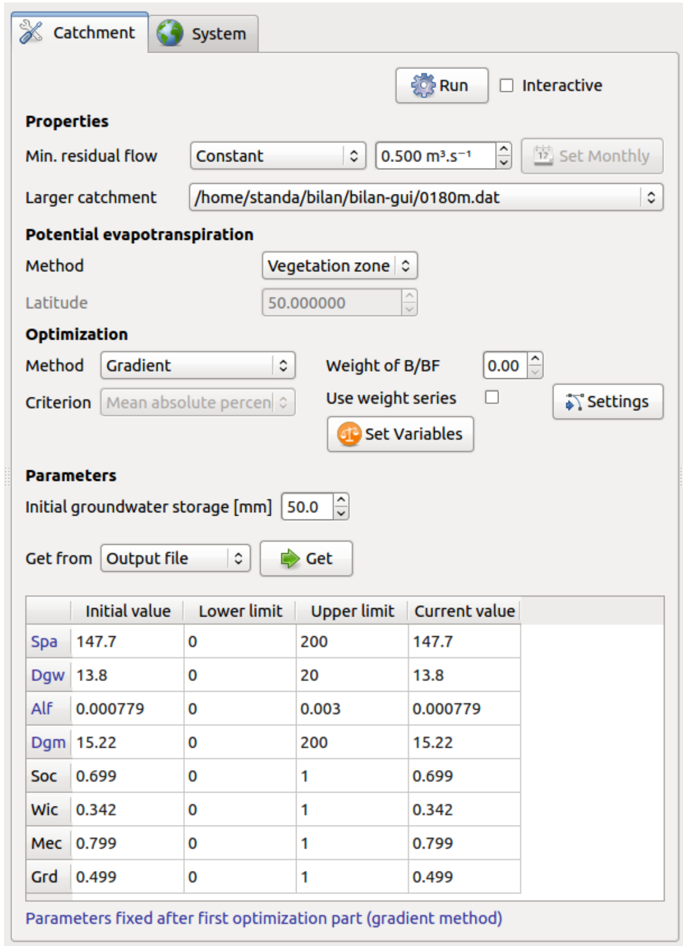


Figure 5: The *Catchment* tab with catchment properties and settings for model run

The *Larger catchment* drop-down allows to choose a catchment whose part the current catchment is. The drop-down contains file names listed on the left. This property is put to use in computations for a system of catchments (see section 7).

4.1 Variables setting

Values of several model variables (currently, observed baseflow and the weights used for optimization) can be set in the *Set Variables* dialog invoked by the *Set Variables* button.

For optimization weights (see also section 5.4) the following options are available:

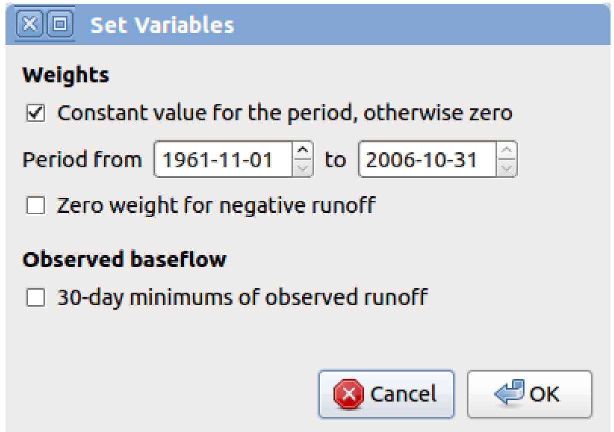


Figure 6: The *Set Variables* dialog for daily data

- ▷ a constant weight value of 1 for a data period specified in the boxes and a zero value for other periods (this can also be set in the time series plot, see section 9),
- ▷ a zero weight value for time steps with negative or missing (NA) values of observed runoff, to suppress the influence of unreliable data.

These two options can be combined.

In case of the daily model, the baseflow time series can be estimated: the baseflow value is computed as the minimum of observed runoff within a 30-day time interval, where the day being estimated is the last day of this interval; a shortened interval is used for days at the beginning of the time series.

All input variables can be transformed in the *Transform Input Data* dialog described in section 3.2.

5 Model settings

Apart from catchment properties, the *Catchment* tab also contains settings needed for running the model. Just as in the case of properties, the current model settings refer to the catchment currently selected from the list of files and may differ from catchment to catchment.

The current model settings for the selected catchment can be saved to the INI file and reused later by using the menu items *Tools* → *Save Catchment Settings* and *Tools* → *Load Catchment Settings* or by the keyboard shortcuts Ctrl+S and Ctrl+L.

5.1 Calculation of potential evapotranspiration

The method is selected in the *Potential evapotranspiration* section. If the *Use input data* item is chosen, the PET data loaded directly from the input or output file will be used.

```
bil.pet
```

Method based on vegetation zones

The potential evapotranspiration is estimated from the saturation deficit by using functions (in the form of tables) derived for individual months and for different bioclimatic zones from empirical graphs given in Gidrometeoizdat (1976). The saturation deficit (in mb) is calculated from air temperature and relative air humidity data. For extreme, improbable or erroneous combinations of these variables, the resulting saturation deficit can either be less than zero or it can exceed the upper calculation limits of the method (nomogram) used to derive the potential evapotranspiration.

For daily data, the PET values are calculated as monthly values divided by 30.

If the saturation deficit is less than zero, the execution will stop and a correction of the data will be required. There is an error display: Physically impossible: negative value of saturation deficit.

If the saturation deficit exceeds the upper calculation limit, the maximum admissible value is substituted.

The bioclimatic zones are as follows: tundra, coniferous forest, mixed forest, deciduous forest and steppe. Each bioclimatic zone is characterised by its characteristic mean air temperature. The model makes use of an interpolation algorithm using the long-term average air temperature in the catchment for reasons of interpolation between the bioclimatic zones, i.e. between the respective tables.

Method based on latitudinal solar radiation and temperature

The potential evapotranspiration is estimated by using a relationship derived by Oudin et al. (2010), employing solar radiation and air temperature and requiring air temperature as the sole input (for details, see Technical reference manual, section 1). Since the value of extraterrestrial solar radiation is calculated for each time step, the catchment latitude (in degrees) has to be specified. As the algorithm used relies on the sunset hour angle, which is not defined for polar day conditions, the latitude entered has to be between the Arctic and the Antarctic Circles (i.e. from -66.5 to 66.5 degrees).

5.2 Initial groundwater storage

The initial conditions in the basin can influence the results of the water balance simulation, particularly at the beginning of the time series. These conditions can be specified by setting the initial groundwater storage value. The default value of initial groundwater storage is 50 mm. It can be changed based either on information on the actual state of the basin in the first time step, or on experience deriving from preliminary simulations.

> The initial groundwater storage in mm is set in the corresponding box in the *Parameters* section.

```
bil.set.optim
```

5.3 Parameters of the model

The free model parameters, which have to be identified for the model to be able to simulate the streamflow generation, are listed in Table 1 for both the daily and the monthly model types.

The optimisation procedure requires initial values of the parameters (relevant for the gradient method only) and their lower and upper limits to be set by the user. The program uses default values, which normally do not have to be changed; however, the values in the table of parameters can be changed to attain an alternative solution.

> The parameter values can be set in the table which shows the initial values and limits. The fourth column of the table is read-only and contains the current optimized values (or the initial values if no optimization has yet been performed). In the table row headers, tooltips appear when a parameter abbreviation is hovered by the mouse.
>
> The initial values of the parameters can also be set by choosing an option from the *Get from* drop-down and pressing the *Get* button:
>
> - The *Output file* option loads the initial parameter values from a file with saved results (old-style formatted output files are also supported). This also sets the number of gradient optimization iterations to zero.
> - The default initial values of parameters and their limits can be restored by the *Default settings* option.

name	daily model type	monthly model type
Spa	capacity of soil moisture storage [mm]	
Dgm	temperature/snow melting factor	
Dgw	–	factor for calculating the quantity of liquid water available on the land surface under winter conditions
Alf	parameter controlling outflow from direct runoff storage (direct runoff)	parameter of rainfall-runoff equation (direct runoff)
Soc	parameter controlling distribution of percolation into direct runoff and groundwater recharge under summer conditions	parameter controlling distribution of percolation into interflow and groundwater recharge under summer conditions
Mec	parameter controlling distribution of percolation into direct runoff and groundwater recharge under conditions of snow melting	parameter controlling distribution of percolation into interflow and groundwater recharge under conditions of snow melting
Wic	–	parameter controlling distribution of percolation into interflow and groundwater recharge under winter conditions
Grd	parameter controlling outflow from groundwater storage (base flow)	

Table 1: Parameters of daily and monthly Bilan model

▷ The initial values of the parameters are replaced by current values (the fourth column) by the *Current values* option.

`bil.set.params`, `bil.set.params.init/lower/upper`, `bil.read.params.file`

5.4 Calibration of the parameters

The parameters of the model are identified (calibrated) by using an optimisation algorithm. The optimisation is aimed at attaining the best fit between the observed and the simulated runoff series. Two optimisation algorithms (the local and global one) are available.

The optimisation settings are located on the *Catchment* tab. The dialog with contextual settings specific to each algorithm is shown by the *Settings* button. A change of optimization method (by the *Method* drop-down) results in new optimization settings with default specific values and their customization for the previous method is lost.

An optimisation run is started by the *Run* button.

If the interactive mode (see section 5.5) is active and an ensemble size greater than 1 is set for DE optimization in the contextual *SCE-UA/DE Optimization Settings* dialog, new ensemble results are added to the list of catchments in each interactive run. These temporary catchments are removed when the dialog is accepted (then new ensemble results are calculated) or cancelled.

`bil.set.optim`

`bil.optimize`, `bil.run`

Local gradient algorithm (binary search)

The calibration of the parameters by using the gradient method is executed in two steps. The procedure under the default settings is described below, however the optimization criterion can be set for each step independently. Apart from the criteria mentioned below, Nash-Sutcliffe efficiency or its logarithmic version can be used. For a detailed description of the criterion functions, see Technical reference manual, section 5.

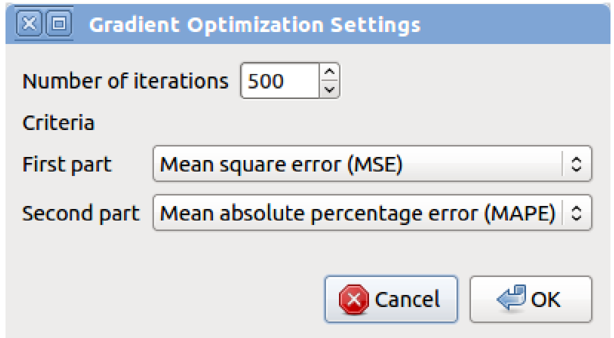


Figure 7: The local gradient optimisation algorithm settings

In the "classical" optimisation procedure, the mean square error (MSE) would normally be used as the optimisation criterion. This criterion suffers from the drawback that its application does not ensure a good fit between the observed and the simulated runoff series in the low-flow parts of the hydrograph. This can substantially be improved by using the sum of relative deviations between the observed and the simulated runoff series (represented by the mean absolute percentage error, MAPE) instead of MSE. However, this criterion frequently deteriorates the fit in terms of the mean runoff and, therefore, an optimisation procedure combining both these criteria has been developed.

In the first step, MSE or the mean absolute error (MAE) is used as the optimisation criterion to calibrate the Dgw (monthly type only), Dgm, Spa and Alf parameters which significantly influence the mean runoff.

> The parameters calibrated in the first step are indicated in blue in the input table of parameters.

The remaining parameters Wic (monthly type only), Mec, Soc and Grd impacting the runoff distribution to its individual components are then calibrated using MAPE. This calibration procedure mostly ensures an acceptable fit in terms of both mean runoff and low-flow runoff consisting predominantly of base flow. The resulting value of the optimization criterion for the second part appears in the output of the model.

Number of iterations The values of the model parameters resulting from the optimisation algorithm can also be influenced by setting the number of iterations to be performed under this algorithm. The default value, derived from practical experience, is 500. Normally, this value need not be changed.

The program can be run also without optimisation of parameters, just by using their initial values. This option can be activated by setting the number of iterations of the optimization procedure to zero.

Extension of parameter limits for initial values For a successful optimization of the gradient type, the initial values of the parameters (set by the user) should not approach the lower or upper limits. If such a situation occurs, the user is notified by the error message: Initial value of parameter 'PARAMETER NAME' is too close to its lower/upper limit.

If the appropriate option is enabled, this error is not reported and the limit of the affected parameter is automatically adjusted before optimization starts. The new value of the lower or upper limit is as close to the original value as possible, whereas it still allows to perform the optimization run.

Global shuffled complex evolution/differential evolution algorithm

The global algorithm employs shuffled complex evolution (SCE-UA) combined with differential evolution for complex evolution (for details, see Technical reference manual, section 4). The user-defined settings of the algorithm are as follows:

- type of differential evolution (best/1/bin, best/2/bin or rand/2/bin),
- number of complexes NC,
- complex size M,
- number of shuffles,
- number of generations,
- crossover parameter CR,
- mutation parameter F,
- mutation parameter K,
- size of ensemble – an ensemble of optimisation runs will be calculated.

Weight of baseflow

By default, the optimisation criterion is calculated from series of observed and simulated runoff. Value of weight for baseflow w_{BF} (between 0 and 1) sets optimisation with respect to difference between observed and simulated baseflow series. The combined criterion is calculated as:

$$\text{crit} = (1 - w_{BF}) \cdot \text{crit}(R, RM) + w_{BF} \cdot \text{crit}(B, BF) \tag{1}$$

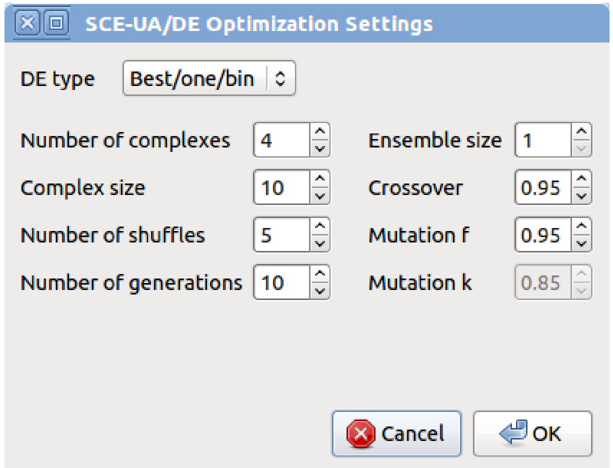


Figure 8: The SCE-UA/DE optimisation algorithm settings

Weights for runoff calibration

Optionally, weights for calibration can be assigned to runoff values in order to emphasize selected parts of the time series or to lower the importance of unreliable data. The weights are loaded as one of input variables. They are considered as relative, so they do not need to sum up to one. Parts of the time series can be excluded from optimization by setting their weights to zero.

If the time series of observed runoff contain any negative or missing (NA) value whose weight is not equal to zero, a warning message that this value affect optimization results is shown before optimization starts.

> To activate the weights for calibration, check the *Use weight series* box. Weight values can be set also by the *Set Variables* dialog (see section 4.1) or in plots (see section 9).

```
bil.set.optim
```

5.5 Interactive run

Model optimization can be run automatically as soon as any parameter or computation setting is changed, allowing to view the results and plots being changed as an interactive response to the user's input. This mode is enabled by checking the *Interactive* box next to the *Run* button on the *Catchment* tab.

The interactive calculations apply for threshold value (impacting the deficit volumes) and for all settings on the *Catchment* tab including the initial parameter values and limits. An interactive run is also possible with settings entered into special dialogs as optimization settings, input data transformation or monthly values dialogs. If such a dialog is cancelled, its settings are restored to the original values and the model is rerun. Interactive changes in these dialogs can involve adding temporary catchments to the list; these are deleted if the dialog is cancelled (for details, see section 3.2 or 5.4). In the *Set Variables* dialog, the interactive run is not supported.

6 Model state initialization

The state variables of the model (reservoir storages and the seasonal mode) can be obtained for the assigned time step and, subsequently (possibly modified), they can be used for model initialization in an arbitrary time step.

This option is not available in GUI.

```
bil.run
```

7 System settings and results

Catchments may be organised so as to form a system. Then, optimization can be performed that will also take the relations between catchments into account.

Catchments within the system have to conform to the same model type. Moreover, their areas have to be set in order to allow for the calculation of the absolute values of flow.

Initially, all catchments on the list are considered to be part of the system, however catchments that do not suit are eventually excluded before optimization. A catchment is connected to another catchment in the system by setting the *Larger catchment* property on the *Catchment* tab.

The structure of the system is shown in a tree form in *System structure* on the *System* tab. When an invalid recurrent structure is chosen or the model types or data periods do not match, a warning message appears instead of the filename identifying the catchment.

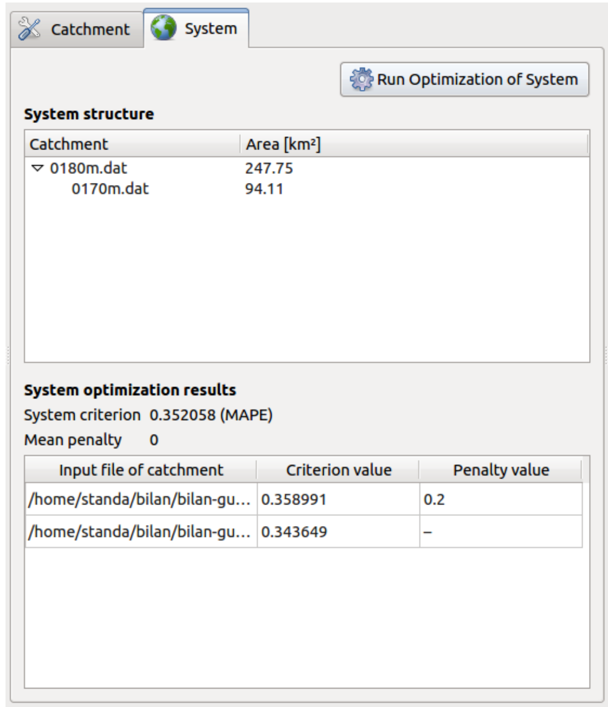


Figure 9: The *System* tab with an optimized system of two catchments

`sbil.new`, `sbil.add.catchs`, `sbil.remove.catchs`

Catchments that do not match the type of the first catchment in the system, as well as catchments that have no connection with it and catchments with no area assigned are removed from the system before optimization. The system cannot be computed if it is of a recurrent structure or contains no catchment.

When a catchment set as larger has actually smaller area assigned, a warning message appears before optimization.

If periods of available data for the catchment and its larger catchment do not match exactly, the common part of them is used when calculating a penalty value.

> For system optimization, the optimization type and settings for the first catchment on the list are applied. This is indicated by the *Also for system* label that is shown atop the optimization settings when the first catchment is selected.
>
> For potential evapotranspiration and deficit volumes, the methods and the threshold values selected for individual catchments are applied.

Optimization is run on pressing the *Run Optimization of System* button. Before starting the optimization, catchments can be excluded based on the rules described above; these rules imply that the first catchment on the list is considered as the first one in the system.

`sbil.set.optim`, `sbil.optimize`

The criterion used when optimizing the parameters of the system of catchments is calculated as the mean value of the criteria for all the catchments. When calculating this mean, the sum of criteria for individual catchments is increased by a penalty, which is function of differences in modelled runoff values between the catchment and its appurtenant larger catchment. Currently, a naive experimental penalty function is applied: it adds the value of 0.1 for any time step showing a negative difference in runoff expressed as the absolute value of flow.

The results of system optimization are shown at the bottom of the *System* tab. At first, the mean criterion and the penalty values are displayed, followed by the table of criteria, the penalties and the number of common time steps for each catchment. The penalty and time step values refer to the relation between the catchment listed on the same line and its appurtenant larger catchment.

Results for individual catchments can be shown on the *Results* and *Plots* tabs when the catchment is selected from the main list on the left-hand side.

If the SCE-UA method is used and the size of the ensemble is greater than 1, the appropriate number of ensemble members is added to the list of catchments.

`sbil.get.optim`, `sbil.get.catch`

8 Results and output

On completion of the optimization, the resulting optimized parameters and mean annual values of all variables are shown on the *Results* tab. The annual values are calculated from monthly characteristics, i.e. from series for complete water years with missing values ignored. If any of the monthly characteristics is of NA value, the corresponding annual will be also NA.

The final value of the optimization criterion is also displayed together with an abbreviation indicating the criterion type (MSE for mean square error, MAE for mean absolute error, MAPE for relative deviations, NS and ln NS for criteria based on Nash-Sutcliffe).

`bil.get.values`, `bil.get.data`, `bil.get.params`

If the SCE-UA method is used and the size of the ensemble is greater than 1, the current file shown in the list on the left represents the first member of the ensemble; results for the other members are stored in catchments denoted as FILE NAME_ensembleNUMBER that have been added to the list.

`bil.get.ens.resul`

The results are written to a file by choosing a type of output in *Type of output series* and pressing the *Save to File* button. If no file extension is entered by the user in the file dialog, the extension .txt is applied automatically.

`bil.write.file`

The output from the model to a file offers three types of series: the daily series (daily type only), the monthly series and the monthly characteristics. The daily or monthly series are time series of variables (both input and output) that have been processed or simulated by the model. In case of the daily type of model the monthly series are aggregated for complete months. If no data are provided or relevant for any given variable or if there is a missing value, NA appears in the time series output.

The monthly characteristics include mean values, minima and maxima derived from the monthly series for each month of the given year and for each variable. They have been calculated from series for complete water years. When calculating the characteristics, missing values are ignored. Hence, the size of dataset represented by the characteristics can differ between months. The characteristics value is NA only if all values for particular month are NAs.

In addition to the selected type of series, the file will include information about the initial date, the optimization criterion value and the optimized parameters. This is shown in an example below. Abbreviations for the variables are listed in Table 2.

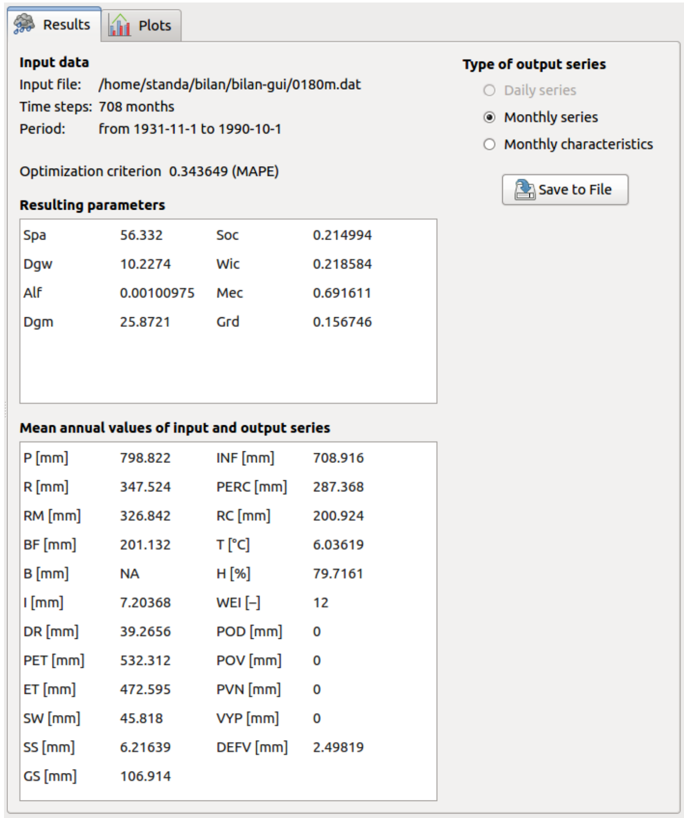


Figure 10: The *Results* tab with calculated values

> Deficit volumes are calculated only in GUI and only if water use variables, catchment area and threshold value are set.

Water balance components (P, R, RM, BF, B, I, DR, PET, ET, INF, PERC, RC, POD, POV, PVN, VYP) are expressed as daily/monthly totals in mm [mm.day^{-1}, mm.month^{-1}]. Components of water storage (SW, SS, GS, DS, DEFV) are expressed as water columns [mm].

> If the *Volume* option is set in the application preferences, the water balance components are expressed as flow [m^3.s^{-1}] (except for the precipitation and evapotranspiration values which remain in mm), and water storage is expressed as volume of water [m^3]. This applies also for the mean annual values shown in the table of results or in plot legend.
>
> The fact that the output file contains the values expressed as flows or volumes is indicated by the text “vol” at the beginning of the line with variable names. When this file is loaded by the *Add From File* button, the measurement units are set automatically (see also section 3.1).

symbol	daily model type	monthly model type
	entry variables	
P	basin precipitation	
T	air temperature	
H	air humidity	
R	runoff (observed)	
B	baseflow (derived)	
WEI	weights for runoff calibration	
POD	groundwater withdrawal	
POV	registered surface water withdrawal	
PVN	unregistered surface water withdrawal	
VYP	water release	
	water balance components – daily or monthly totals	
PET	potential evapotranspiration	
ET	basin evaporation	
INF	infiltration into the soil	
PERC	percolation through the soil layer	
RC	recharge of groundwater storage	
I	–	interflow – amount of water exceeding soil moisture capacity
DR	direct runoff	
BF	base flow (simulated)	
RM	total runoff (simulated)	
	components of water storage	
SS	snow water storage	
SW	soil moisture (water storage in zone of aeration)	
GS	groundwater storage	
DS	direct runoff storage	–
DEFV	deficit volumes	
	other variables (not listed in output file)	
POT	–	water from rainfall or snowmelt (quantity limited by the air temperature)
AKT	–	water in the form of snow available for melting and subsequent infiltration
MELT	snow melting	–
RDS	inflow to direct runoff storage	–

Table 2: Abbreviations of variables of the daily and monthly Bilan models

The data series form a matrix. Each row of the matrix includes data for one particular day or month (e.g. 1st of November 1977 or November 1971), while individual variables form columns of the matrix.

In the output of monthly characteristics, each variable forms a matrix. Rows of the matrices contain data for individual months of the year (e.g. November), while the columns of the matrices contain month identification, the minimum value, mean value and maximum value.

The data from the output file can be loaded again as described in section 3.1.

Example of monthly time series output:

```
Initial
1931-11-1

Spa     56.332
Dgw     10.2274
Alf     0.00100975
Dgm     25.8721
Soc     0.214994
Wic     0.218584
Mec     0.691611
Grd     0.156746

OK      0.343649

P       R       RM      BF      B       I       DR      PET     ET      SW      SS      GS      INF     PERC
RC      T       H       WEI
5.957   17.658  7.87312 7.83729 NA      0       0.0358  9.19836 9.10486 53.1483 0       42.1627 5.92117 0
0       1.454   84      1
63.117  20.115  15.8074 6.60883 NA      9.1986  0       4.57095 4.57095 56.332  13.2797 68.438  45.2664 42.0827
32.8841 -3.574  87      1
...     ...     ...     ...
```

Example of monthly characteristics output:

```
Initial
1931-11-1

Spa     56.332
Dgw     10.2274
Alf     0.00100975
Dgm     25.8721
Soc     0.214994
Wic     0.218584
Mec     0.691611
Grd     0.156746

OK      0.343649

month   min     mean    max
P
```

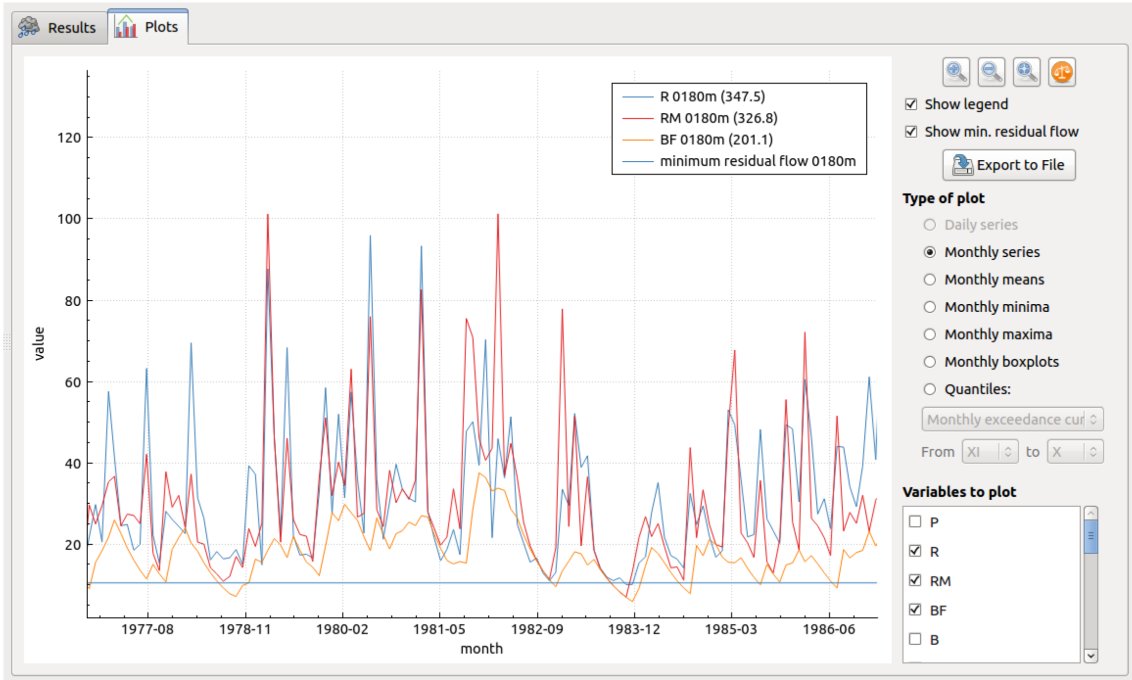


Figure 11: The *Plots* tab with a zoomed plot of monthly time series

```
11      5.957   58.8554 136.212
12      2.098   59.9924 151.283
1       16.76   54.5892 191.02
2       6.177   48.9115 148.65
3       4.053   45.6621 97.628
4       12.035  49.7532 106.812
5       14.209  77.7308 156.832
6       38.282  92.3674 234.564
7       13.201  101.426 254.736
8       24.409  90.5929 213.021
9       3.723   62.7433 194.695
10      0.663   56.198  180.72

R
11      8.129   22.6984 60.226
...     ...     ...     ...
```

9 Plots of results

Results can be visualized on several types of plots, including time series of given variables and their statistical characteristics. All plotting options can be set and the resulting plot is displayed on the *Plots* tab.

The plot shows the values of the currently selected catchment. To show values for a catchment permanently, the *Show always in plots* box below the catchment list needs to be checked when the required catchment is selected. This feature is supported in all plot types.

Variables to be plotted are checked in the *Variables to plot* list. A tooltip with variable name appears when a variable abbreviation is hovered by the mouse.

Time series are shown by selecting the *Daily series* or *Monthly series* option. In case of monthly series for daily data, aggregated monthly values are used. Four basic types of monthly characteristics are available: means, minima, maxima and boxplots where boxes consist of the first and third quartile and median and whiskers range from minimum to maximum of the series.

The X axis scale of plots of time series can be controlled by the *Zoom in* and *Zoom out* icons. By pressing one of these icons an interactive mode of scaling is enabled, then the scale can be changed also by scrolling using the mouse wheel. By pressing the *Show all* icon, the complete time series are shown and the interactive mode is disabled.

In time series plots, grey background indicates periods for which zero optimization weights are set. The period of constant non-zero weight can be changed interactively by pressing the *Set period of constant weight* icon (which enables the appropriate mode) and clicking into the plot. After that, weights are set for a period between the two time steps marked by clicking, in the same way as when using the option in the *Set Variables* dialog (see section 4.1).

For more specific analyses, quantile plots of monthly data can be displayed. *Exceedance curves* show the probability of exceedance of variable values, where the empirical probability p is estimated as

$$p = \frac{m - 0.3}{n + 0.4} \tag{2}$$

where n is the number of values (time steps) and m is the sequential order of a value ranging from 1 to n.

Gumbel plots, focusing on extreme values of the monthly series (minima or maxima), show the probability in the form of Gumbel reduced variate G which is calculated as

$$G = -\ln(-\ln p) \tag{3}$$

The *Q-Q plot* is a scatter plot of corresponding quantiles of observed runoff R and simulated runoff RM.

When creating boxplots and quantile plots, missing values are ignored. Therefore number of values used may be different for different months (as it is for calculation of characteristics described in section 8).

A range of months to be considered can be specified by the drop-down lists *From…to* for all types of quantile plots. This makes it possible to investigate seasonal characteristics of the time series.

In all types of plots, NA values are represented by missing parts. Therefore, if a daily or monthly series contain NAs, unfilled points are drawn together with lines (otherwise values surrounded by NAs would not be visible).

The plot legend is shown by checking the *Show legend* box. The legend items consist of variable abbreviation, file name of catchment and annual mean of the variable. No legend is provided for Q-Q plots.

If a threshold value for a catchment is defined, the *Show threshold value* checkbox allows to show the constant threshold value in all types of plots or monthly values in time series and characteristics plots.

Any plot can be saved as a PNG image or PDF document by using the *Export to File* button. If no file extension is entered by user in the file dialog, the extension .png or .pdf is applied automatically.

10 Model modifications

Apart from the daily and monthly default model types, there currently are two models available which have a modified structure or modified functions. These modified versions are considered to be experimental and not suitable for productive use.

The modifications are not available in GUI.

```
bil.new
```

10.1 Only one parameter dividing percolation

In this modification, only one parameter called Soc is used for division of percolation instead of three (monthly) or two (daily version) parameters Soc, Mec, Wic. The replaced parameters still appear in model outputs although they do not influence the results in any way.

10.2 Arbitrary time step

This modification allows to use an arbitrary time step given as the number of days. The modified version differs only in the PET calculation; otherwise, algorithms for the original daily or monthly model are applied. Therefore, e.g. the daily structure for weekly or even monthly data can be used. It is worth noting that the parameter values for the modified time step are of different scale than those resulting from the original model, and cannot be compared with them.

10.3 Criterion variables and weights

In this modification, user can set a custom criterion function by defining an arbitrary number of pairs of observed and modelled variables and the corresponding weight value. It is a generalization of the case when runoff and baseflow time-series are used for optimization (described in section 5.4).

```
bil.set.critvars
```

Technical reference manual

In the technical manual, the internal structure of the model and its individual algorithms are described.

1 Description of the Oudin method for PET estimation

The potential evapotranspiration can be estimated by the method derived by Oudin et al. (2010). The relation for the PET value on a given day i requires only the air temperature T as input:

$$PET(i) = \begin{cases} \dfrac{0.408 R_e (T(i)+5)}{100} & \text{if } T(i)+5>0 \\ 0 & \text{if } T(i)+5 \leq 0 \end{cases} \tag{4}$$

where R_e denotes extraterrestrial radiation [MJ,m^{-2}d^{-1}] that can be obtained as described in Allen et al. (1998):

$$R_e(i) = \frac{24 \cdot 60}{\pi} G_{SC}\, d_r [\omega_s \sin\varphi \sin\delta + \cos\varphi \cos\delta \sin\omega_s] \tag{5}$$

where G_{SC} is a solar constant (0.082 MJ,m^{-2}min^{-1}), d_r is the inverse relative distance Earth-Sun

$$d_r = 1 + 0.033 \cos\left(\frac{2\pi}{365} J\right) \tag{6}$$

where J is number of the day in the year, δ is the solar declination (angular distance to the equator) [rad]

$$\delta = 0.409 \sin\left(\frac{2\pi}{365} J - 1.39\right) \tag{7}$$

ω_s is the sunset hour angle [rad]

$$\omega_s = \arccos[-\tan\varphi \tan\delta] \tag{8}$$

and φ is the catchment latitude [rad].

2 Description of the daily model type

2.1 Type of regime

The application of some of the model algorithms is dependent on the conditions prevailing on the particular day. Using mean daily air temperature, the model distinguishes between summer conditions and winter conditions. The summer conditions are assumed if the temperature

$$T(i) \geq 0 \tag{9}$$

If there is a snow cover in the basin, a snow melting algorithm is used instead of the summer algorithm.

2.2 Components of total runoff

The model simulates the total runoff $RM(i)$ as the sum of two components:

$$RM(i) = DR(i) + BF(i) \tag{10}$$

where $DR(i)$ and $BF(i)$ are direct runoff and base flow, respectively.

The direct runoff $DR(i)$ represents a fast response of the catchment and is constituted by the outflow from direct runoff storage.

The base flow $BF(i)$, whose retention time in the basin is longer than that of the direct runoff component, is constituted by the outflow from groundwater storage.

If water use variables are included, the total runoff is decreased by surface water withdrawal (registered $POV(i)$ and unregistered $PVN(i)$) and increased by the amount of water release $VYP(i)$, as long as the runoff value is positive:

$$RM(i) = \max(DR(i) + BF(i) - POV(i) - PVN(i) + VYP(i), 0) \tag{11}$$

2.3 Evaporation and water balance in the soil under summer conditions

If infiltration, which equals to precipitation, equals or exceeds potential evapotranspiration

$$INF(i) \geq PET(i) \tag{12}$$

the basin evaporation is equal to the potential evapotranspiration

$$ET(i) = PET(i) \tag{13}$$

the excess water amounting $INF(i) - PET(i)$ is available to feed the soil moisture

$$SW(i) = SW(i-1) + INF(i) - PET(i), \tag{14}$$

and if capacity of the soil moisture storage is exceeded

$$SW(i) > Spa, \tag{15}$$

the remaining water percolates downwards

$$PERC(i) = SW(i) - Spa \tag{16}$$

and the soil moisture storage $SW(i)$ is equal to the parameter Spa.

If the potential evapotranspiration exceeds the precipitation (infiltration), the basin evaporation is supplied from the soil moisture storage being depleted

$$SW(i) = SW(i-1) \cdot e^{\frac{INF(i)-PET(i)}{Spa}} \tag{17}$$

where e is the base of natural logarithms.

The basin evaporation is equal to the sum of the precipitation and the soil moisture depletion

$$E(i) = INF(i) + SW(i-1) - SW(i) \tag{18}$$

and no water is available for percolation.

2.4 Evaporation and water balance on the land surface under winter conditions

If the sum of precipitation and water storage in the snow cover on a given day exceeds potential evapotranspiration, it is assumed that the basin evaporation is equal to the potential evapotranspiration,

$$ET(i) = PET(i), \qquad \text{for} \quad SS(i-1) + P(i) \geq PET(i) \tag{19}$$

else the basin evaporation is calculated as the sum of precipitation and snow water storage

$$ET(i) = SS(i-1) + P(i), \qquad \text{for} \quad SS(i-1) + P(i) < PET(i) \tag{20}$$

The quantity of the remaining water stored in the form of snow is

$$SS(i) = \max(SS(i-1) + P(i) - PET(i), 0) \tag{21}$$

where $SS(i-1)$ is water storage in the snow cover on the day $i-1$.

The infiltration is supposed to be zero

$$INF(i) = 0 \tag{22}$$

2.5 Evaporation and water balance on the land surface under snow melting conditions

If the precipitation on the given day exceeds the potential evapotranspiration, it is assumed that the basin evaporation is equal to the potential evapotranspiration

$$ET(i) = PET(i) \qquad \text{for} \quad P(i) > PET(i) \tag{23}$$

otherwise the basin evaporation equals the precipitation rate

$$ET(i) = P(i) \qquad \text{for} \quad P(i) \leq PET(i) \tag{24}$$

The quantity of water produced by melting (available for infiltration) is limited by the heat capacity of the air to melt the snow cover on the given day, which is expressed as

$$MELT(i) = T(i) \cdot Dgm \tag{25}$$

where $T(i)$ is the mean air temperature on the day i and Dgm is a parameter expressing the rate of snowmelt as a function of air temperature.

The remaining snow water storage is then computed as

$$SS(i) = \max(SS(i-1) - MELT(i), 0) \tag{26}$$

Water from snow melting is assumed to infiltrate, eventually together with precipitation excess:

$$INF(i) = MELT(i) \qquad \text{for} \quad P(i) \leq PET(i) \tag{27}$$

$$ET(i) = MELT(i) + P(i) - PET(i) \qquad \text{for} \quad P(i) > PET(i) \tag{28}$$

2.6 Water balance in soil under snow melting conditions

Water calculated as infiltration $INF(i)$ supplies the soil moisture (or zone of aeration), which is assumed to have its capacity given by the parameter Spa. If the soil capacity is exceeded, the excess water, $PERC(i)$, percolates downwards to feed inflow to direct runoff storage and recharge of groundwater storage. In other words, if the sum of soil water storage from the preceding day, $SW(i-1)$, and infiltration in the given day, $INF(i)$, exceeds the Spa parameter, it holds that

$$PERC(i) = SW(i-1) + INF(i) - Spa \tag{29}$$

$$SW(i) = Spa \tag{30}$$

Otherwise

$$PERC(i) = 0 \tag{31}$$

$$SW(i) = SW(i-1) + INF(i) \tag{32}$$

2.7 Distribution of percolation into direct runoff and groundwater recharge

Percolation $PERC(i)$ is divided into runoff component $RDS(i)$ that flows to direct runoff storage and recharge $RC(i)$ that replenishes the groundwater storage

$$RDS(i) = c \cdot PERC(i)^2 \tag{33}$$

$$RC(i) = \max(PERC(i) - RDS(i), 0) \tag{34}$$

In the above equations, the parameter Mec is substituted for c under snow melting conditions and by the parameter Soc in summertime.

2.8 Water balance in direct runoff storage and direct runoff

Direct runoff storage $DS(i)$ on the day i is calculated as the sum of the storage from the preceding day and inflow $RDS(i)$:

$$DS(i) = RDS(i) + (1 - Alf) \cdot DS(i - 1) \tag{35}$$

The direct runoff $DR(i)$ represented by the outflow from the storage is controlled by the parameter Alf

$$DR(i) = Alf \cdot DS(i) \tag{36}$$

2.9 Water balance in groundwater and base flow

Groundwater storage $GS(i)$ on the day i is calculated as the sum of the storage from the preceding day and the recharge $RC(i)$. The base flow represented by the outflow from the groundwater is proportional to its storage at the beginning of the given day and is controlled by the parameter Grd

$$BF(i) = Grd \cdot GS(i - 1) \tag{37}$$

Therefore, the groundwater storage at the end of the day is

$$GS(i) = RC(i) + (1 - Grd) \cdot GS(i-1) \tag{38}$$

If the water use variables are included, the groundwater storage is decreased by groundwater withdrawal $POD(i)$ as long as the value of storage remains positive:

$$GS(i) = \max(RC(i) + (1 - Grd) \cdot GS(i-1) - POD(i), 0) \tag{39}$$

3 Description of the monthly model type

3.1 Type of regime

Some of the model algorithms applied depend on the conditions prevailing during the given month.. Using mean monthly air temperature, the model distinguishes between summer conditions and winter conditions. Summer conditions are assumed if the temperature

$$T(i) \geq 0 \tag{40}$$

If there is a snow cover on the basin, a snow melting algorithm is used instead of the summer algorithm.

3.2 Components of total runoff

The model simulates the total runoff $RM(i)$ as the sum of three components:

$$RM(i) = DR(i) + I(i) + BF(i) \tag{41}$$

where $DR(i)$, $I(i)$ and $BF(i)$ are direct runoff, interflow and base flow, respectively.

The $DR(i)$ component of the total runoff includes summer surface runoff and that part of interflow which, together with the surface runoff, flows so rapidly that it neither affects water balance in the soil nor is significantly available for evaporation. The summertime direct runoff is caused by heavy rainfall.

Irrespective of the season, the interflow $I(i)$ results from water balance as excess water in the zone of aeration. This runoff component is assumed also to include surface runoff if it occurs in winter or during the period of snow melting.

The base flow $BF(i)$, whose retention time in the basin is longer than that of the other runoff components, is constituted by the outflow from groundwater storage.

If the water use variables are included, the total runoff is decreased by surface water withdrawal (registered $POV(i)$ and unregistered $PVN(i)$) and increased by the amount of water release $VYP(i)$, as long as the runoff remains positive:

$$RM(i) = \max(DR(i) + I(i) + BF(i) - POV(i) - PVN(i) + VYP(i), 0) \tag{42}$$

3.3 Formation of direct runoff under summer conditions

Direct runoff occurring during the summer season due to heavy rainfall is calculated as

$$DR(i) = Alf \cdot P(i)^2 \cdot \frac{SW(i-1)}{Spa} \tag{43}$$

where Alf is a parameter of the quadratic rainfall-runoff relationship between direct runoff and rainfall, $P(i)$ is precipitation in the month i, $SW(i-1)$ is soil moisture in the month $i-1$, and Spa is a parameter expressing soil moisture capacity.

The precipitation reduced by the direct runoff

$$INF(i) = P(i) - DR(i) \tag{44}$$

becomes a component of water balance in the zone of aeration.

3.4 Evaporation and water balance in the soil under summer conditions

If precipitation reduced by direct runoff, $INF(i)$, calculated by Equation 44 equals or exceeds the potential evapotranspiration

$$INF(i) \geq PET(i), \tag{45}$$

the basin evaporation is equal to the potential evapotranspiration

$$ET(i) = PET(i), \tag{46}$$

the excess water amounting $INF(i) - PET(i)$ is available to feed the soil moisture

$$SW(i) = SW(i-1) + INF(i) - PET(i), \tag{47}$$

and if the capacity of soil moisture storage is exceeded

$$SW(i) > Spa, \tag{48}$$

the remaining water percolates downwards

$$PERC(i) = SW(i) - Spa \tag{49}$$

and the soil moisture storage $SW(i)$ is equal to the parameter Spa.

If potential evapotranspiration exceeds the precipitation reduced by direct runoff, the basin evaporation is supplied from the soil moisture storage being depleted:

$$SW(i) = SW(i-1) \cdot e^{\frac{INF(i)-PET(i)}{Spa}} \tag{50}$$

where e is the base of natural logarithms.

The basin evaporation is equal to the sum of the reduced precipitation and the soil moisture depletion

$$ET(i) = INF(i) + SW(i-1) - SW(i) \tag{51}$$

and no water is available for percolation.

3.5 Evaporation and water balance on the land surface under winter conditions and during the period of snow melting

If the sum of precipitation and water storage in the snow cover in the given month exceeds potential evapotranspiration, it is assumed that the basin evaporation is equal to the potential evapotranspiration

$$ET(i) = PET(i). \tag{52}$$

The quantity of the remaining water, which is potentially available for infiltration (disposable water in the form of snow), is

$$AKT(i) = SS(i-1) + P(i) - PET(i) \tag{53}$$

where $SS(i-1)$ is water storage in the snow cover in the month $i-1$.

However, the potential quantity of water available for infiltration is limited by the heat capacity of the air to melt the snow cover during the given month, which is, under snow melting conditions, expressed as

$$POT(i) = T(i) \cdot Dgm + P(i) \tag{54}$$

where $T(i)$ is the mean air temperature in the month i and Dgm is a parameter expressing the rate of snowmelt as a function of air temperature.

Under winter conditions, a part of precipitation is assumed to be constituted by rainfall or a part of the existing snow cover will succumb to melting, particularly if the monthly air temperature exceeds a certain value set implicitly as $Tepk = -8$°C.

Again, the amount of water that will be available in liquid form is determined by using the air temperature

$$POT(i) = (T(i) - Tepk) \cdot Dgw \tag{55}$$

and is controlled by the parameter Dgw.

If the mean monthly air temperature is below the value specified as $Tepk$, the water balance on the land surface is described by the following equation

$$SS(i) = SS(i-1) + P(i) - PET(i) \tag{56}$$

and $INF(i) = 0$, which means that no water infiltrates into the soil, and the difference between precipitation and potential evapotranspiration is added to the snow water storage.

Under both winter and snow melting conditions, if the disposable water $AKT(i)$ exceeds the limit $POT(i)$, the $AKT(i)$ is split into a part that infiltrates, $INF(i)$, and a part that remains on the land surface as snow cover. Then it holds that

$$INF(i) = POT(i) \tag{57}$$

$$SS(i) = AKT(i) - INF(i). \tag{58}$$

If the limit $POT(i)$ exceeds the quantity of disposable water, this water is fully available for infiltration

$$INF(i) = AKT(i) \tag{59}$$

and water storage in the snow cover is exhausted.

The $AKT(i)$ value can exceptionally be negative, when the sum of precipitation and water storage in the snow cover in the given month is lower than potential evapotranspiration, and thus

$$INF(i) = 0, \tag{60}$$

$$SS(i) = 0, \tag{61}$$

$$ET(i) = P(i) + SS(i-1). \tag{62}$$

3.6 Water balance in soil under winter and snow melting conditions

Water calculated as infiltration $INF(i)$ supplies the soil moisture (or zone of aeration), which is assumed to have its capacity given by the parameter Spa. If the soil capacity is exceeded, the excess water, $PERC(i)$, percolates downwards to feed recharge of groundwater storage and interflow. In other words, if the sum of soil water storage from the preceding month, $SW(i-1)$, and infiltration in the given month, $INF(i)$, exceeds the parameter Spa, it holds that

$$PERC(i) = SW(i-1) + INF(i) - Spa, \tag{63}$$

$$SW(i) = Spa. \tag{64}$$

Otherwise

$$PERC(i) = 0, \tag{65}$$

$$SW(i) = SW(i-1) + INF(i). \tag{66}$$

3.7 Distribution of percolation into interflow and groundwater recharge

Percolation $PERC(i)$ is divided into runoff component $I(i)$ that reaches the stream channel in the given month and recharge $RC(i)$ that replenishes the groundwater storage

$$I(i) = c \cdot PERC(i), \tag{67}$$

$$RC(i) = (1 - c) \cdot PERC(i). \tag{68}$$

In the above equations, the parameter Wic is substituted for c under winter conditions, the parameter Mec for snow melting and the parameter Soc in summertime.

3.8 Water balance in groundwater and base flow

Groundwater storage $GS(i)$ in the month i is calculated as the sum of the storage in the preceding month and the recharge $RC(i)$. The base flow represented by the outflow from groundwater is proportional to its storage at the beginning of the given month and is controlled by the parameter Grd:

$$BF(i) = Grd \cdot GS(i - 1). \tag{69}$$

Therefore, the groundwater storage at the end of the month is

$$GS(i) = RC(i) + (1 - Grd) \cdot GS(i - 1). \tag{70}$$

If the water use variables are included, the groundwater storage is decreased by groundwater withdrawal $POD(i)$ as long as the value of storage remains positive:

$$GS(i) = \max(RC(i) + (1 - Grd) \cdot GS(i - 1) - POD(i), 0) \tag{71}$$

4 Description of the SCE-UA optimization algorithm

The global algorithm employed combines the SCE-UA (shuffled complex evolution – The University of Arizona) method as described in Duan et al. (1994) and the differential evolution (DE) method by Storn and Price (1997) which is used for complex evolution.

The algorithm steps are as follows:

1. A population of a given number NP of parameter sets (points) is generated by using Latin Hypercube sampling with specified upper and lower limits for the parameters.

2. The parameter sets are sorted by their criterion values.

3. The parameter sets are divided into NC complexes, each complex containing M sets. The best criterion value set p_1 is assigned to the first complex C_1, the second criterion value p_2 to the complex C_2 and so on, i.e. the complex C_1 contains sets p_1, p_{NC+1} up to $p_{(M-1)\cdot NC+1}$.

4. Complexes are evolved by the differential evolution method.

5. A new population is created from the evolved complexes.

6. The algorithm stops on reaching the maximum number of iterations. Otherwise, it continues to step 2.

The differential evolution algorithm for complex evolution (step 4) is described as follows:

1. Mutation is performed on the best parameter set of the whole population p_B or a random parameter set p_{r5} by using differences between parameter sets randomly chosen from the complex. Three types of mutation are available:

$$mp_{i,G+1} = p_{B,G} + F\left(p_{r1,G} - p_{r2,G}\right) \tag{72}$$

$$mp_{i,G+1} = p_{B,G} + F\left(p_{r1,G} - p_{r2,G}\right) + K\left(p_{r3,G} - p_{r4,G}\right) \tag{73}$$

$$mp_{i,G+1} = p_{r5,G} + F\left(p_{r1,G} - p_{r2,G}\right) + K\left(p_{r3,G} - p_{r4,G}\right) \tag{74}$$

   where G and $G+1$ denote a generation of parents and offsprings, F and K are mutation control parameters and $r1, r2 \ldots r5$ are random indexes of parameter set within the complex.

2. If the probability of crossover for a given parameter par exceeds the value of the crossover control parameter CR, then the parameter value of the mutated set is assigned to offspring:

$$p_{i,G+1}[par] = mp_{i,G+1}[par] \tag{75}$$

   Apart from the probabilistic crossover depending on CR, the mutated value for one parameter of random index i is always assigned to offspring.

   Otherwise, or if the mutated parameter has exceeded its limits and if such values are bound to be rejected, the parent parameter value is assigned to offspring:

$$p_{i,G+1}[par] = p_{i,G}[par] \tag{76}$$

3. Selection: If the criterion value for the set of offspring parameters is better than that for parent, the new parameter set is assigned to offspring, else the old parental parameters are retained for the offspring.

The algorithm can be allocated a name composed of the type of parameter set to be mutated (best/rand), the number of differences between parameters used for mutation (1/2) and the type of crossover (binomial). Therefore, Equation 72 represents the best/1/bin type, while Equation 73 is best/2/bin and Equation 74 is rand/2/bin.

5 Description of criterion functions

The mean square error (MSE) is the mean of squared deviations between the observed and the simulated runoff series:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(RM(i) - R(i))^2 \tag{77}$$

The mean absolute error (MAE) is calculated as the mean of absolute deviations between the observed and the simulated runoff series, where "absolute" means that negative deviations are converted to positive values:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|RM(i) - R(i)| \tag{78}$$

The mean absolute percentage error (MAPE) represents the mean of relative deviations where "relative" means that each deviation is divided by the observed value:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|RM(i) - R(i)|}{R(i)} \tag{79}$$

The definition implies that MAPE cannot cope with zero values of runoff.

Nash-Sutcliffe efficiency (NS) or logarithmic Nash-Sutcliffe (LNNS) can be also used:

$$NS = 1 - \frac{\sum_{i=1}^{n}(RM(i) - R(i))^2}{\sum_{i=1}^{n}(R(i) - \bar{R})^2} \tag{80}$$

$$LNNS = 1 - \frac{\sum_{i=1}^{n} (\ln RM(i) - \ln R(i))^2}{\sum_{i=1}^{n} (\ln R(i) - \overline{\ln R})^2} \tag{81}$$

where $\bar{R}$ is the mean observed runoff and $\overline{\ln R}$ is the mean of the logarithmized runoff time series:

$$\overline{\ln R} = \frac{1}{n} \sum_{i=1}^{n} \ln R(i) \tag{82}$$

Bibliography

Allen, Richard G., Luis S. Pereira, Dirk Raes, and Martin Smith. 1998. *Crop evapotranspiration – Guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56.* Rome: FAO – Food; Agriculture Organization of the United Nations.

Duan, Qingyun, Soroosh Sorooshian, and Vijai K. Gupta. 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology* 158: 265–284. doi:10.1016/0022-1694(94)90057-4.

Gidrometeoizdat. 1976. *Rekomendatsii po roschotu ispareniia s poverhnosti suchi.* St. Peterburg: Gidrometeoizdat.

Oudin, Ludovic, Lætitia Moulin, Hocine Bendjoudi, and Pierre Ribstein. 2010. Estimating potential evapotranspiration without continuous daily data: possible errors and impact on water balance simulations. *Hydrological Sciences Journal* 55: 209. doi:10.1080/02626660903546118.

Storn, Rainer, and Kenneth Price. 1997. Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *J. of Global Optimization* 11: 341–359. doi:10.1023/A:1008202821328.