SINCLAIR KNIGHT MERZ

SKM

Department of Agriculture, Fisheries and Forestry

Using baseflow for monitoring stream condition and groundwater and surface water resource condition change

DECISION SUPPORT TOOL – BASEJUMPER

USER MANUAL

- Final
- 22 June 2007

Department of Agriculture, Fisheries and Forestry

Using baseflow for monitoring stream condition and groundwater and surface water resource condition change

DECISION SUPPORT TOOL – BASEJUMPER

USER MANUAL

- Final
- 22 June 2007

Sinclair Knight Merz
ABN 37 001 024 095
590 Orrong Road, Armadale 3143
PO Box 2500
Malvern VIC 3144 Australia
Tel: +61 3 9248 3100
Fax: +61 3 9248 3400
Web: www.skmconsulting.com

Contents

Document history and status

Revision	Date issued	Reviewed by	Approved by	Date approved	Revision type
Draft	21/2/2007	B.Neal	B.Neal	1/4/07	draft
Final	22/6/2007	B. Neal	B.Neal		

Distribution of copies

Revision	Copy no	Quantity	Issued to
Draft	1	1	Steering Committee
Final	1	1	Steering Committee
	2	1	BaseJumper download website (www.skmconsulting.com, search term “BaseJumper”)

Printed:	27 June 2007
Last saved:	26 June 2007 04:56 PM
File name:	I:\WCMS\Projects\WC03396\Deliverables\Milestone 5 - Decision Support Tool\R02rem_UserManual.doc
Author:	Rachel Murphy
Project manager:	Brad Neal
Name of organisation:	Department of Agriculture, Fisheries and Forestry
Name of project:	Using baseflow for monitoring stream condition and groundwater and surface water resource condition change
Name of document:	Decision Support Tool User Manual
Document version:	Final
Project number:	WC03396

1. Introduction

1.1 What is this document?

These notes are designed to aid an uninitiated user in running BaseJumper. This document gives a basic description of the operation of BaseJumper, and provides instructions on how to input data, estimate the baseflow component of streamflow, and identify trends in the separated baseflow data. It is assumed that the user has some experience in using data management and manipulation software packages, such as Excel.

1.2 What is BaseJumper?

BaseJumper is a rapid assessment tool that can be used to provide water resource managers with insights about surface water and groundwater interaction. It provides users with information on the potential impact that anthropogenic influences independent of climate (such as groundwater pumping, land use change, etc) have on the rate of change of baseflow contribution to streamflow. This tool is expected to be used as a pre-cursor to more detailed modelling or field investigations in areas where baseflow contributions to streamflow are identified as changing over time by this tool. The use of this tool forms part of an integrated management approach to identify and manage rivers with declining baseflow.

1.3 What can BaseJumper do?

BaseJumper allows the user to input daily streamflow data and then perform baseflow separation using a digital recursive filter. Trend analysis of this data can then be undertaken within the tool to determine the changes in baseflow contribution to streams independent of climate variability.

1.4 What can't BaseJumper do?

BaseJumper has been developed for application at locations where a baseflow signal is identifiable. However, baseflow signals may not be readily identifiable in some locations and hence the tool may not be suitable for application in these areas. A decision tree is incorporated into the tool (refer to Section 2.1), which provides users with the ability to determine whether the tool can be used to identify trends in baseflow at their site.

While BaseJumper can identify trends in baseflow, the reasons for the trends can not always be easily inferred. In many locations, the interaction of a range of landuse changes, such as farm dam development, diversions and groundwater pumping, are likely to impact on the trends identified. Additionally, stream reaches may be gaining or losing, or alternate between the two, either spatially or temporally. Consequently, it is important for users to be aware of the local conditions in order to have a reasonable appreciation of the processes occurring within their study catchment, and use this information to understand how baseflow contributions to streamflow may be affected by these processes.

1.5 What is provided?

BaseJumper is supplied as an executable file with associated sample files. The executable file allows for the application of the BaseJumper software, as outlined in this document. The sample files are provided to allow the user to test the software using pre-prepared data.

1.6 How do I install BaseJumper on my computer?

BaseJumper can be installed by downloading and saving the setup zip file to a suitable location. The zip file contains the program executable, help files, user manual and sample data. The contents should be extracted and the BaseJumper.exe file can be opened directly.

2. Basic application of BaseJumper

There are three key steps to using BaseJumper:

1) Import data;
2) Baseflow separation; and
3) Trend analysis.

These steps should be undertaken in the above order, as described in the following sections of this document.

2.1 Site selection

A suitable site must be selected before the tool can be applied. Factors to be considered in the selection of a site include:

- Availability of data (refer to Section 2.2 for data requirements);
- Knowledge of local conditions: it is important that the user be informed on the local catchment conditions, particularly those that may influence the baseflow contributions to streamflow. This should include essential information such as:
  - Catchment characteristics, for information on the current and historical land uses.
  - Regional hydrogeology, to gain an appreciation of the connectivity between groundwater and surface water resources. In particular, it is important to be aware of gaining and losing reaches, and the implications this has on baseflow contribution to the stream in the area of interest. Baseflow is not present in losing streams.
  - Streamflow management, which includes flow conditions, diversions, discharges, streamflow regulating structures.
    - Major diversions – diversions for consumptive use such as irrigation channels, urban diversions, etc. These diversions can decrease low flows and hence appear to reduce estimates of baseflow. Allowances can be accurately made for those diversions where they are metered.
    - Return flows – Water can be returned to rivers from sewage treatment plants or from industry. Power stations in particular often discharge cooling tower water to streams. This will increase low flows and appear similar to baseflow.
    - Flow regulation from upstream reservoirs: reservoirs that release outflows that are different to inflows will produce a low flow signal that can be misinterpreted as baseflow at downstream flow gauges. Neal et.al. (2004) adopted a criterion that not more than 10% of the catchment should be upstream of flow regulating structures

when selecting streamflow gauges for regional baseflow assessment, which should be regarded as an upper limit for local investigations.

Streams with diversions, discharges and flow regulating structures require additional data processing prior to application of BaseJumper, and it is recommended that users avoid these river reaches if possible. The steps described in Section 2 are only relevant to reaches without these external influences. To apply BaseJumper to a reach with diversions, discharges or flow regulating structures, the reader is encouraged to consolidate skills in applying the tool using a simpler site before progressing to the steps described in Section 0.

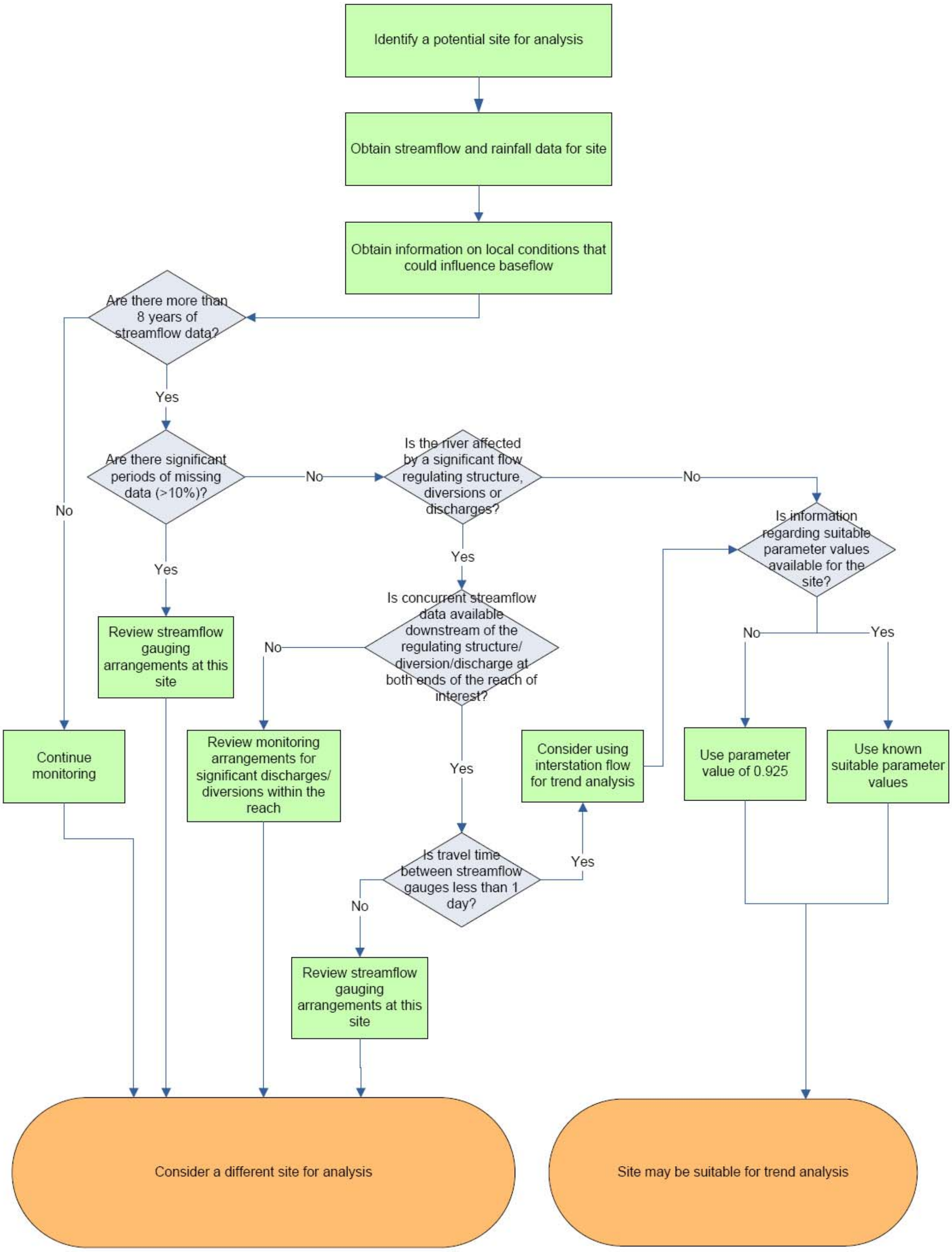
Additional desirable information about the local catchment includes:

- Groundwater use and management (regardless of whether the impacts of groundwater extractions are to be investigated through the use of BaseJumper) to understand the direct impacts to the groundwater resources in the region.
- Catchment farm dams: high concentrations of catchment farm dams could also influence baseflow but only where the dams are located on-stream or where they interact with groundwater. Many off-stream dams are clay lined specifically to avoid interaction with groundwater.
- Urbanisation – In urban areas, activities such as excess garden or sports field watering can increase low flows during summer that appear similar to baseflow in streamflow data (Daamen et al, 2006).
- River evaporation and evapotranspiration – Evaporation from the river surface and plant water uptake will generally be a negligible influence on streamflows for catchments less than 1000 km^2. For larger catchments, baseflow expressed at an upstream location may be reduced at the streamflow gauging station because of these reach losses, particularly during summer low flow conditions when baseflow is most evident.

- Availability of information on suitable regional digital recursive parameter values. While this is not essential, it ensures that the results are most accurate given the local conditions. In the absence of regional digital recursive filter parameter information, the user can use the information in Section 2.5.2 to attempt to identify a suitable parameter value. Alternatively, a value of 0.925 is widely accepted as a default parameter value based on the work of Nathan and McMahon (1990)

These factors are summarised in Figure 2-1. This flow chart should be consulted before a study site is finalised.

- **Figure 2-1 Decision tree for determining site suitability**



2.2 Data Requirements

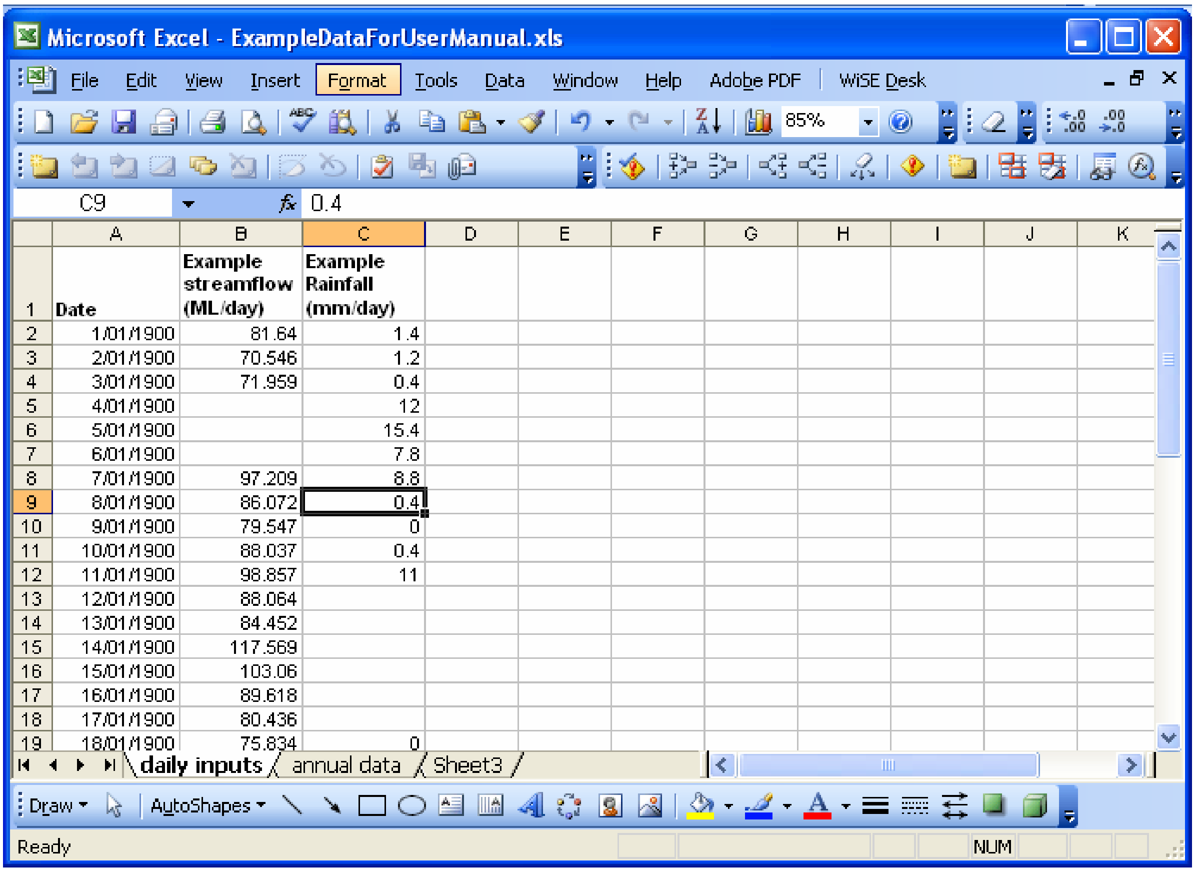
Several forms of data are necessary to run BaseJumper, including:

- Daily streamflow data: A data set of at least 8 complete years is recommended. If possible, the quality codes associated with each data point should be investigated, and periods of low quality data should be removed from the record (data quality codes should be available from the organisation responsible for the collection of streamflow data in your area). Low quality data includes that obtained from unreliable measurement, extrapolated data, data lost to natural causes, data recorded during an equipment malfunction, and data not collected. Periods of missing data should have blank entries for the streamflow on the appropriate dates (do not replace missing data with zero or negative flow values), and should account for less than 10% of the total data record. Daily streamflow should be in units of ML/day.
- Optional variables for regression analysis: Changes in the baseflow contributions to streamflow may be attributed to a variety of activities. The selection of other optional regression analysis variables requires the user to have some understanding of the processes and activities occurring within the study catchment. Data sets could include climate, groundwater volumes pumped, percentage tree cover (for tree plantations) and other variables that could reflect land use changes. While up to three variables may be incorporated into the regression analysis, in some situations it may not be necessary to consider this many variables. The number and nature of the variables to be included in the trend analysis will depend on the characteristics of the study catchment.
  - Where rainfall data is to be included, this should be added as a daily time series at the same time as streamflow data. A dataset that is concurrent with the streamflow data is required. The rainfall gauge selected for use should be located within the catchment of the chosen streamflow gauge and should be representative of the local catchment conditions upstream of the flow gauge being analysed. The Bureau of Meteorology's Water Resources Station Catalogue (http://www.bom.gov.au/hydro/wrsc?searchAOI=2) can be used to search for a rainfall gauge close to the streamflow gauge location. Where a rainfall gauge with data available over a concurrent period is not available within the streamflow gauge catchment, data should be sourced from the closest gauge with a suitable data record. As described for the streamflow data, periods of missing data should be left blank. The rainfall data should be in units of mm/day.
  - With the exception of rainfall, other variables must be input into BaseJumper on an annual basis, where the user defines the start and end months of the annual analysis. This additional data must be concurrent with the period of streamflow and rainfall data. Data is required to be manipulated into an annual format external to BaseJumper. Missing or low quality data should be replaced with blank cells, or infilled using an appropriate method.

SKM

It is suggested that data is stored within an external data software program, such as Excel. This allows the user to manage and manipulate the data as required prior to input into BaseJumper. Figure 2-2 provides an example screen shot of the daily data required for input to BaseJumper. Figure 2-3 provides an example screen shot of the annual data required for input to BaseJumper.

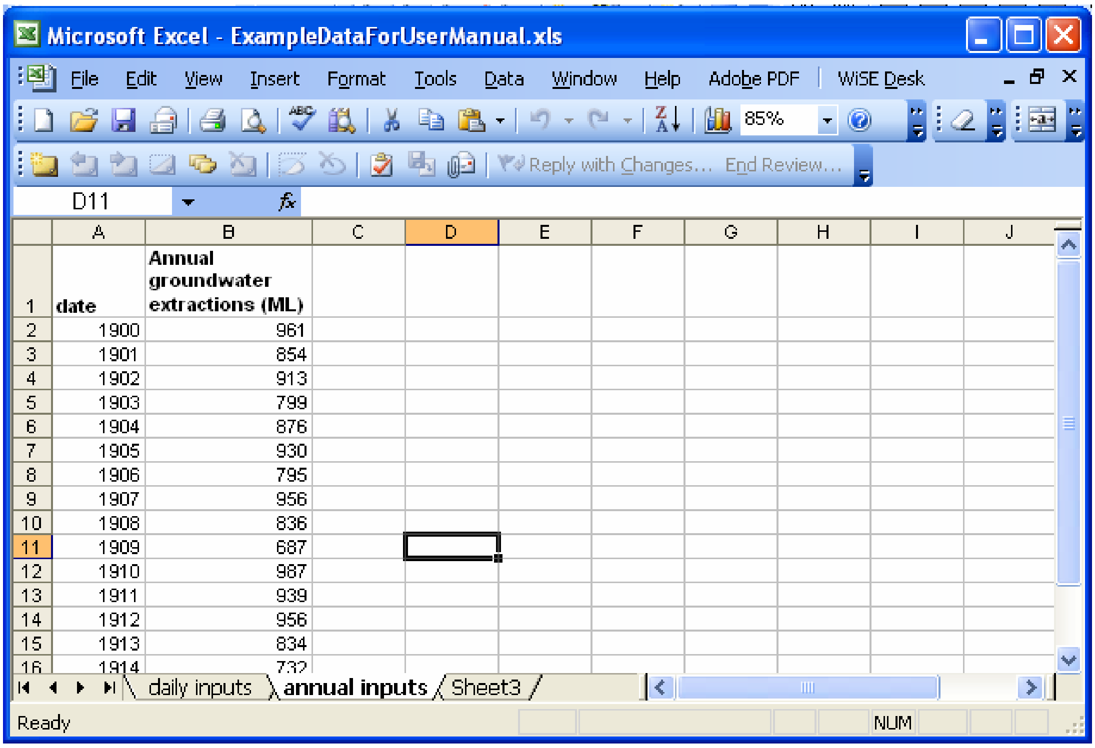
- **Figure 2-2 Example daily input data. Missing or low quality data is replaced with blank cells**



Date	Example streamflow (ML/day)	Example Rainfall (mm/day)
1/01/1900	81.64	1.4
2/01/1900	70.546	1.2
3/01/1900	71.959	0.4
4/01/1900		12
5/01/1900		15.4
6/01/1900		7.8
7/01/1900	97.209	8.8
8/01/1900	86.072	0.4
9/01/1900	79.547	0
10/01/1900	88.037	0.4
11/01/1900	98.857	11
12/01/1900	88.064	
13/01/1900	84.452	
14/01/1900	117.569	
15/01/1900	103.06	
16/01/1900	89.618	
17/01/1900	80.436	
18/01/1900	75.834	0

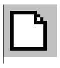
SKM

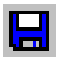
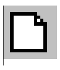
- **Figure 2-3 Example annual input data**

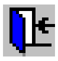


2.3 BaseJumper File functionalities

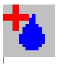
The BaseJumper File menu provides the user with various options regarding data management.

File>New or the  icon allows the user to create a new BaseJumper file for analysis. File>Open or the  icon allows the user to access previously saved data. By selecting this option, the user can navigate to the appropriate location and open an existing BaseJumper file (a BaseJumper file is of the type *.bj).

File>Save and the  icon will be available to the user once analysis is commenced within BaseJumper. When File>New or the  icon is initially selected, the user is forced to save the output text file to a suitable location with a sensible file name to provide documentation on the data inputs. This also ensures an auditable path with information on data preparation being retained. If required, the outputs of this text file can be later transferred to Excel for further data analysis. A dialogue box will indicate any problems with the save, or will inform the user that the data has been saved as requested. At key steps throughout the BaseJumper analysis, the user will be asked whether they would like to save their data again. Additionally, the user can elect to save the data again at any stage through the analysis process, and it is recommended that data be saved regularly. The save function will append the latest task to the existing text file. If a number of regressions are undertaken, only the latest one will be recorded in the output file when saved. The user should undertake a new analysis if several regressions results are required.

File>Exit or the  icon closes the BaseJumper program.

2.4 Step 1: Importing data

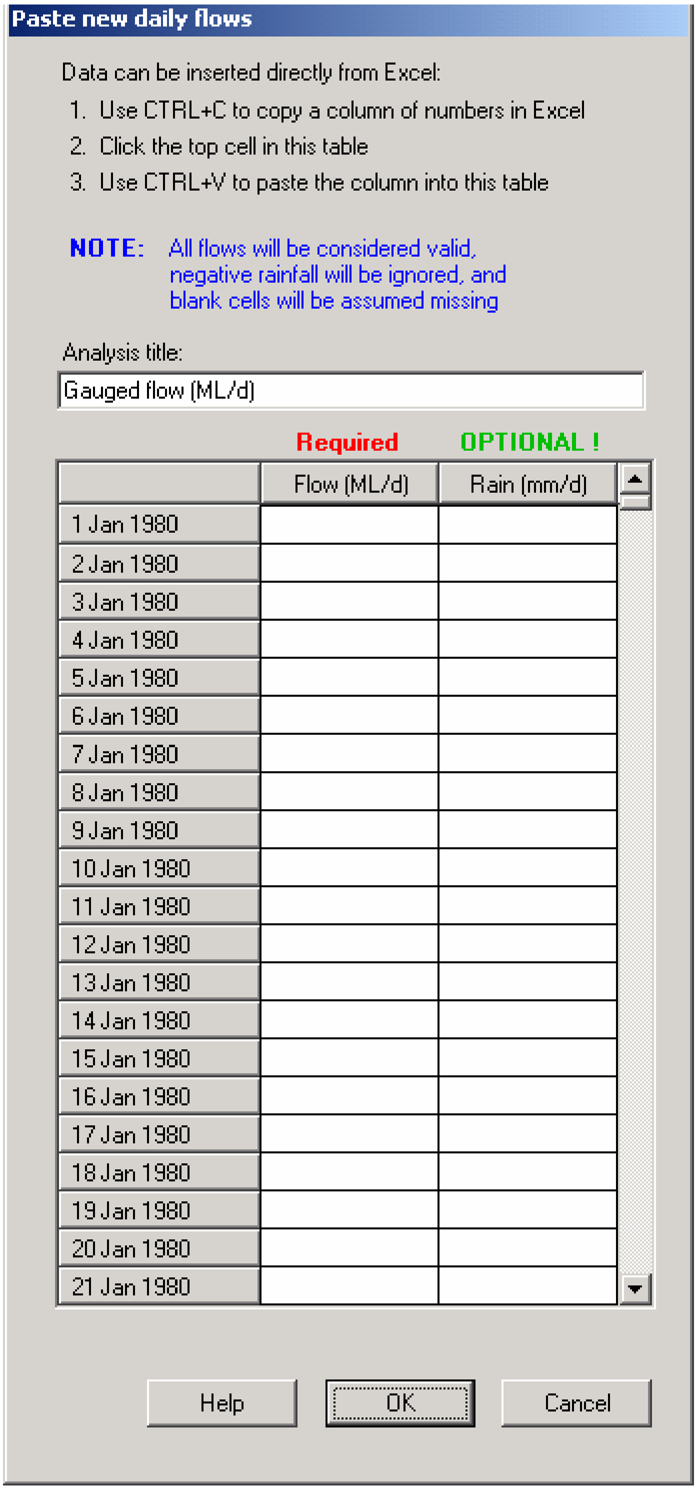
The Analysis>Add New Daily Flows menu option or the  icon allows a user to start a new analysis. Daily streamflow data is the first data to be provided to BaseJumper. The user will be prompted to provide the starting date of the data. A worksheet style display will open with dates down the left hand column, commencing from the start date provided by the user. A column will be available for streamflow data. By selecting the first cell of the flow data column and ensuring the worksheet is at the start of the record, streamflow data from Excel can be directly copied and pasted into the tool (Select the relevant streamflow data in Excel, use *Control+C* to copy this data, and *Control+V* to paste it into BaseJumper). Data corresponding to each day will be shown down the time series. Negative streamflow data will be considered valid. This is relevant when interstation flows are being analysed (refer to Section 0).

An additional column in the worksheet is available for daily rainfall data. This can be copied and pasted as described for the streamflow data. It is not essential to input rainfall data, as both the baseflow separation and trend analysis can be undertaken without this information. Negative rainfall data will be ignored (assumed missing).

Figure 2-4 shows BaseJumper daily input worksheet with the example streamflow and rainfall data. A description of the analysis can be described in the Analysis Title text box.

SKM

▪ **Figure 2-4 BaseJumper daily input worksheet**



Once the streamflow and rainfall data has been imported, the user will be provided with some information about the data sets, including details on the length of the record and periods of missing data.

The user will be prompted to update the saved text file at this point. If the user would like to save the work completed thus far, the output file will be annotated with the input streamflow and rainfall time series.

The Analysis>Review /edit daily flows menu option (or icon) will now be available to the user. This can be selected at any stage throughout the analysis process to review the daily streamflow and rainfall data within the tool.

2.5 Step 2: Baseflow Separation

2.5.1 Baseflow Separation Theory

The digital filter is based on the theory described by Nathan and McMcMahon (1990) and uses the Lyne and Hollick filter:

$$q_f(i) = \alpha q_f(i-1) + \frac{[q(i) - q(i-1)] \times (1+\alpha)}{2} \quad \text{for} \quad q_f(i) \geq 0$$

▪ **Equation 2-1**

Where:

$q_f(i)$ is the filtered quickflow response for the i^{th} sampling instant

$q(i)$ is the original streamflow for the i^{th} sampling instant

α is the filter parameter that enables the shape of the separation to be altered.

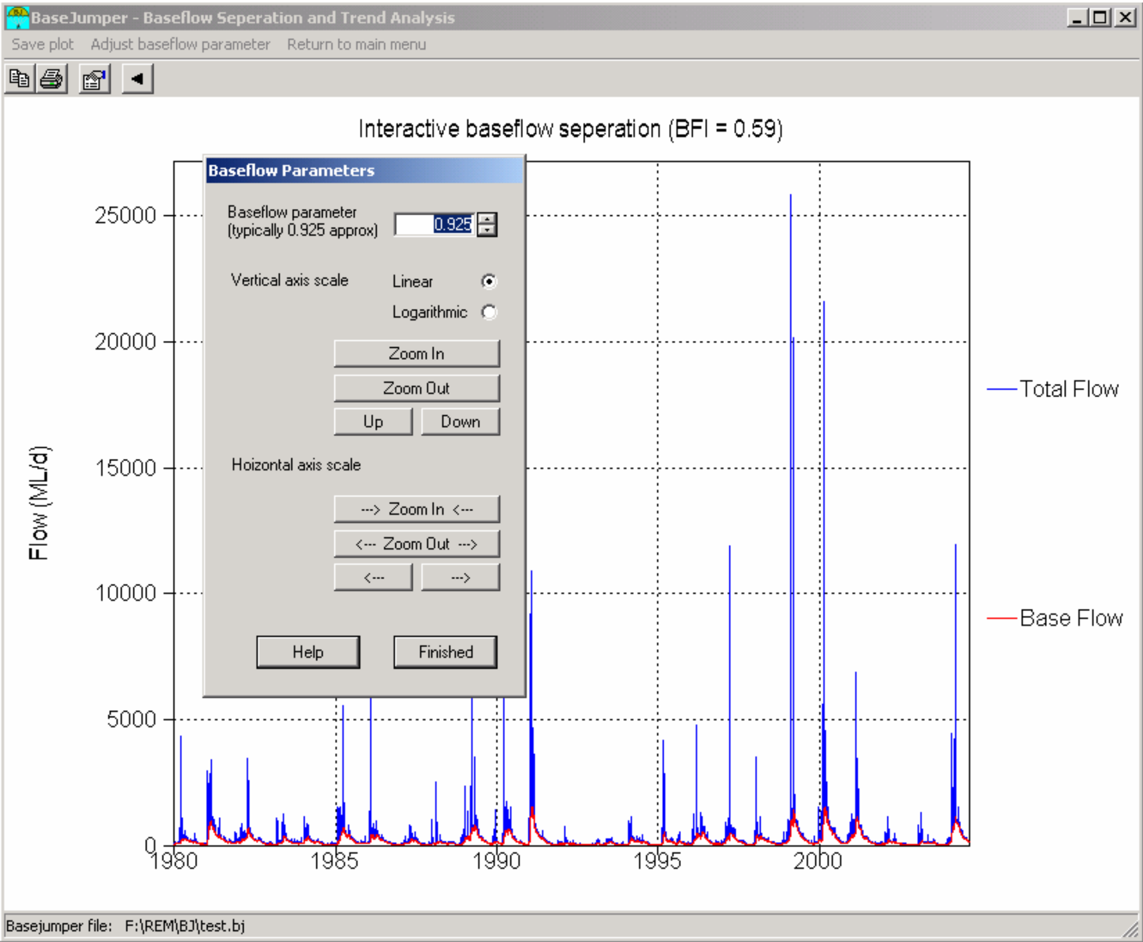
After applying this equation, the baseflow ($q_b(i)$) is equal to $q(i)$ - $q_f(i)$. If $q_f(i)$ is less than zero, then $q_b(i)$ is set equal to $q(i)$. This filter is run in three passes – the first and third passes are "forward" passes using the equation above directly, whereas the second pass is a "backward" pass using i+1 rather than i-1 in the equation. In the first pass, $q(i)$ is the computed streamflow, in the second pass it is the computed baseflow from the first pass, and in the third pass it is the computed baseflow from the second pass. These passes act to smooth the data.

This separation technique is an automated procedure that produces a baseflow timeseries that may subsequently be manually adjusted for more detailed investigations.

2.5.2 Baseflow Separation using BaseJumper

The Analysis>Baseflow separation menu function (or icon) starts the baseflow separation process. The user is presented with a plot of the streamflow data and the separated baseflow. An interactive parameter window allows the user to modify the digital recursive filter parameter and zoom in on the data. This window is shown in Figure 2-5.

Figure 2-5 Interactive baseflow separation window



The most appropriate digital recursive filter parameter to apply is commonly 0.925. However, this depends on the location and conditions of the site being investigated. While, theoretically, the parameter value can be any positive value less than 1, research has shown that values between 0.90 and 0.99 ensure a plausible baseflow time series is generated. Values less than 0.90 tend to cause the baseflow hydrograph to rise too steeply, while higher parameter values produce a baseflow hydrograph that is too flat and has its peak preceding the total hydrograph peak. Hence, in order to ensure the separated baseflow is plausible, BaseJumper restricts the user to values of 0.90-0.99.

When selecting a parameter value for a particular site, it is recommended that the user tests the results with a parameter value of 0.925. If the separated baseflow does not appear plausible, the user can examine the shape of the separated baseflow and compare it to the streamflow hydrograph to observe features such as:

- Inclusion of surface runoff in the baseflow hydrograph – this is indicated by a steep rise in the baseflow hydrograph at the commencement of streamflow events;
- Timing of the peaks in the baseflow hydrograph – the baseflow hydrogrpah should peak after the total hydrgraph due to the storage-routing effect of the sub-surface storages;
- Steepness of the peaks in the baseflow hydrograph – low parameter values result in high baseflow hydrograph peaks, while high parameter values result in lower baseflow hydrograph peaks. Consider the plausibility of these peaks;
- Baseflow recession behaviour in log domain – the baseflow hydrograph will most likely follow an exponential decay function (a master recession curve). In the log domain, the baseflow recession should appear linear; and
- General baseflow hydrograph behaviour in high and low flow periods – consider the extent of interflow or quickflow in the baseflow hydrograph, and the magnitude of the baseflow hydrograph peaks during these events.

As an example, the following plots show how these factors could be used to determine a plausible separated baseflow time series.

Figure 2-6 presents a one year period of example streamflow data. The separated baseflow data is also shown for a variety of digital recursive filter parameters. This figure has been annotated with comments on the appearance of the baseflow hydrograph. Examples of the characteristics listed above are provided. Figure 2-7 shows the same data set in the logarithmic domain. Important features of the baseflow hydrograph are again identified. It is recommended that the user consider such characteristics for each baseflow separation undertaken using BaseJumper, to ensure the resulting baseflow time series is plausible. Precisely identifying baseflow from streamflow data introduces a degree of subjectivity and the user should consider analysing a range of hydrographs using different filter parameter values if there is uncertainty about the nature of the baseflow hydrograph.

Figure 2-6 Example streamflow and baseflow components (linear domain)

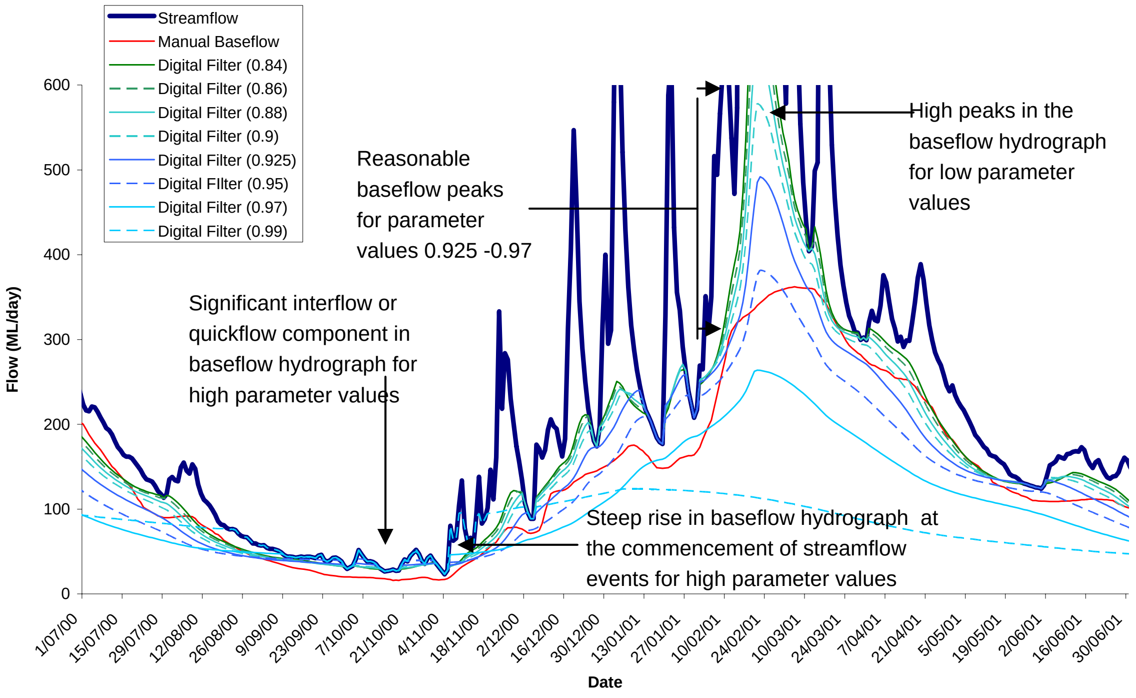
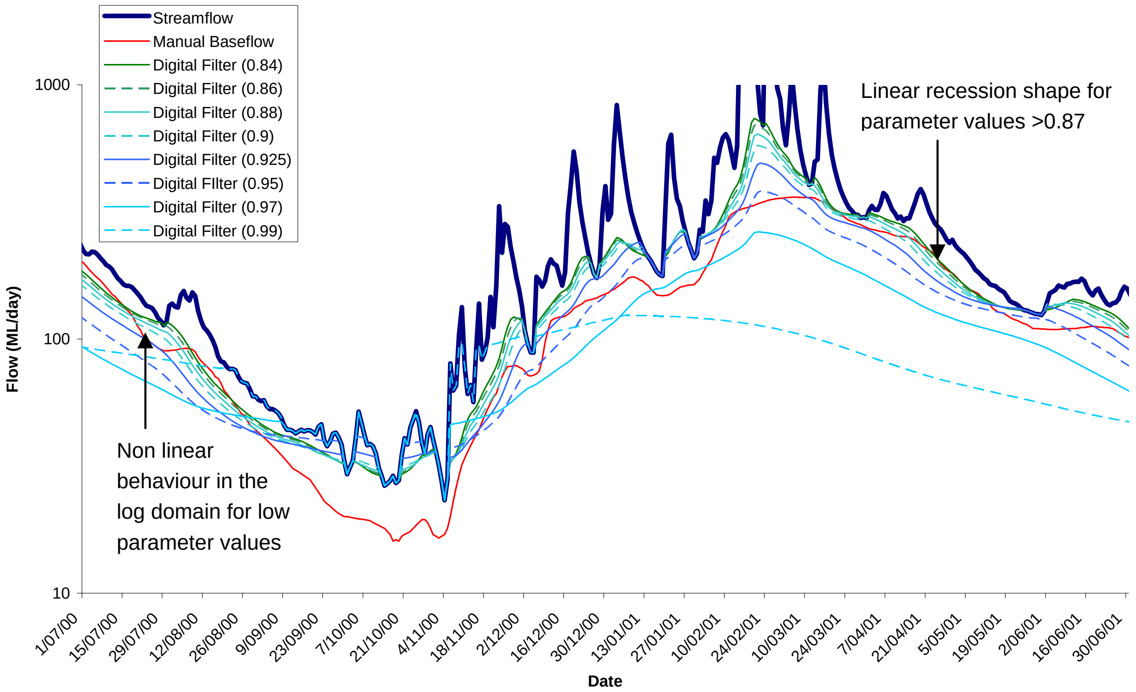


Figure 2-7 Example streamflow and baseflow components (logorithmic domain)



An interactive plot of the total streamflow and baseflow is available to assist the user in examining the separated baseflow. Many of the features listed above can be observed in this manner. This plot is automatically opened when Analysis>Baseflow separation (or icon) is selected. The user is able to modify the digital recursive filter parameter using the toggle button to observe the impact the parameter value has on the separated baseflow time series. The default setup of the plot uses a parameter value of 0.925 and shows the full time series on an arithmetic scale. The user is able to zoom and scroll through the entire time series using the relevant arrows on the interactive parameter window. A logarithmic scale is also available. The plot will automatically update with any changes made in this interactive window.

Once the baseflow time series has been finalised, the user selects Finished on the interactive parameter window. The user can save the results of the baseflow separation at this stage. The Save Plot menu option or the icon saves the baseflow timeseries to the clipboard for pasting into Word. By selecting the Return to Main Menu (icon), the user is again prompted to save the work so far. If necessary, the user can return to the interactive parameter window by choosing the Adjust Baseflow Parameter menu option (or the icon). BaseJumper automatically retains the information from the last baseflow separation, should the user want to return to this analysis without having to update the parameter value.

It is possible to use BaseJumper for just the baseflow separation process. If the user has no need to complete the trend analysis process, BaseJumper can be exited once the separation process is complete.

2.6 Step 3: Data aggregation

Daily streamflow (and possibly rainfall data), which were input to the tool through the process described in Section 2.4, and the baseflow timeseries separated through the process described in Section 2.5 need to be aggregated to annual data sets for the trend analysis. BaseJumper completes this task for the user, based on information provided about the period for aggregation.

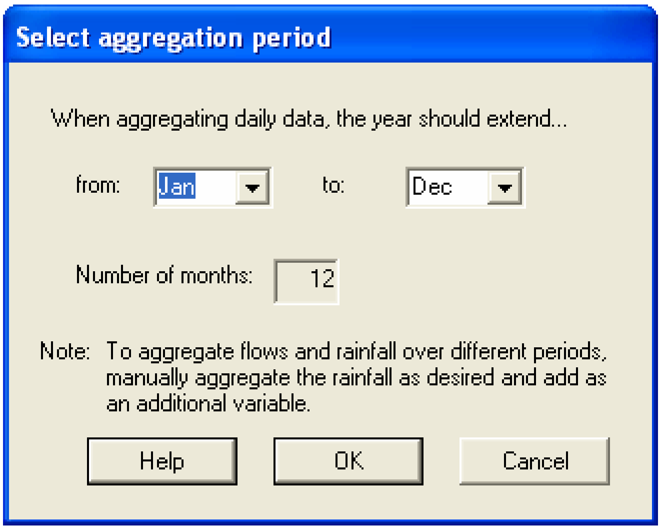
The Analysis>Select Aggregation Period menu option opens a dialogue box (Figure 2-8) which allows the user to indicate the start and end months of the aggregation period. The user is able to specify the period over which data will be aggregated. Typically, a 12 month period will be applied. This may be based on a calendar year (January – December) or could reflect water management policies (for example, a water year extending from July – June). Alternatively, a shorter period could be used for aggregation. This may be appropriate for the analysis of sites in highly seasonal environments. For example, the user may like to consider the dry season impacts.

The user should select the relevant months in the data aggregation dialogue box. BaseJumper will then sum the flow, baseflow and rainfall over these months for each year on record, to generate an annual time series of dry season data.

The dialogue box informs the user of the duration of the period for aggregation. The user should check that the intended duration matches that indicated by BaseJumper. Subsequent annual data provided to BaseJumper (refer to Section 2.7.2) should be aggregated over a consistent period.

Should it be necessary to aggregate the rainfall and baseflow over different periods (for example, in some environments it may be relevant to consider the rainfall during one season and the baseflow during another), the user should manually aggregate the rainfall and add it as an additional variable in the trend analysis step.

- **Figure 2-8 Select aggregation period dialogue box**



2.7 Step 4: Trend Analysis

2.7.1 Trend Analysis Theory

Research during the development of the tool determined that a linear model was able to adequately reproduce the results obtained using a more rigorous model. Hence, the trend analysis functionality in BaseJumper uses a linear regression model.

Multiple linear regression is a statistical technique that allows one dependent variable to be predicted from a number of independent variables. A regression equation has the form:

$$Y = a_o + a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

Equation 2-2

Where:

Y is the dependent variable

a is a coefficient

X is an independent variable

n refers to the number of independent variables selected

In order to determine the coefficients (a_0, a_1, etc), a data set is required for which all of the dependent and independent variables are known. The coefficients in the equation are determined using a method of least squares. This method selects the coefficients so that the difference between the observed (recorded) value and the value predicted by the linear regression equation is minimised.

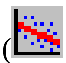
In order to draw statistical inferences from the trend analysis, three basic assumptions need to be satisfied, namely that model residuals (difference between the observed value and the value predicted by the linear regression equation):

1) Are normally distributed;
2) Have constant variance; and
3) Are serially independent (are not correlated).

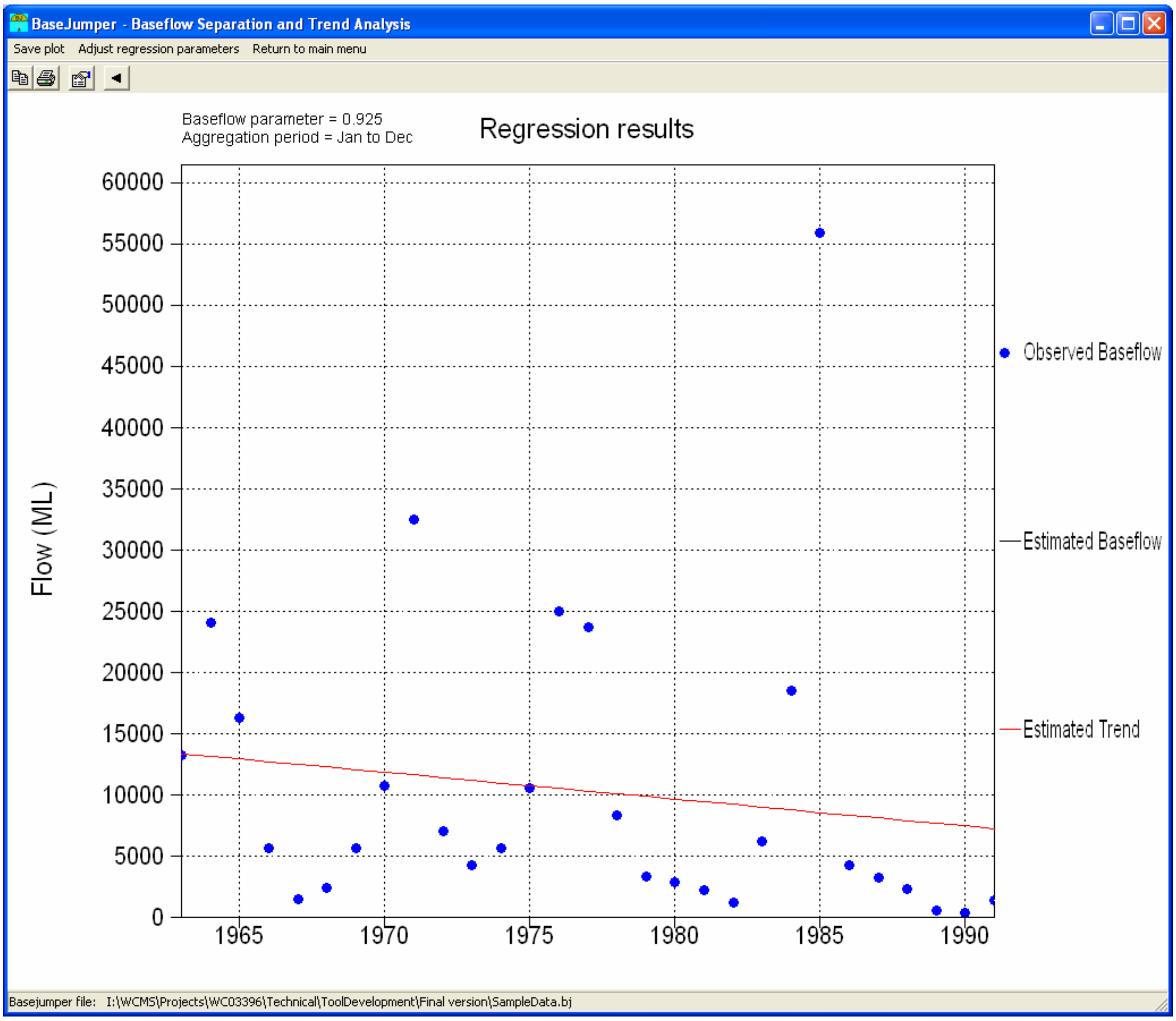
These assumptions need to be checked before a regression equation is adopted.

Outliers are excluded in the statistical analysis using an objective criterion based on Cook's statistic (Cook, 1977 and Cook, 1979). Cook's statistic evaluates the sensitivity of the model to outliers by considering the influence of each data point on the coefficients of the fit of the data. A value of Cook's statistic greater than 2 is commonly used as the basis for the identification of outliers. The use of Cook's statistic enables outliers to be identified without detailed examination of individual data pairs, which is considered impractical for many large scale analyses.

2.7.2 Trend Analysis using BaseJumper

Simple linear regression analysis is undertaken to determine any trends in baseflow over time. The Analysis>Baseflow regression menu option ( icon) commences the trend analysis operation. Data previously input into the tool (daily streamflow and possibly rainfall data) will have been automatically aggregated to an annual time step based on the time period specified by the user in the data aggregation step. BaseJumper produces a simple plot displaying the average change in baseflow over time (Figure 2-9). The plot also specifies the parameter used in the baseflow separation and the data aggregation period used.

■ **Figure 2-9 Average change in baseflow over time**



A baseflow regression dialogue box displays some statistical information about the relationship, including details of the trend significance, model fit and flow statistics. The parameter box also allows the user to include or ignore outlying data points. Cook's statistic is used to determine whether input data points are considered outliers. The user can choose to include outliers by selecting the tick box on the regression parameter screen.

The user should review this plot to observe the change in baseflow over time, and to identify whether further trend analysis is required.

If rainfall data was provided to BaseJumper, this can be added to this trend analysis by ticking the '*Include?*' rainfall option. This will provide a linear trend analysis that is independent of climate, that is:

$$Baseflow = a_0 + a_1[Rainfall] + a_2[time]$$

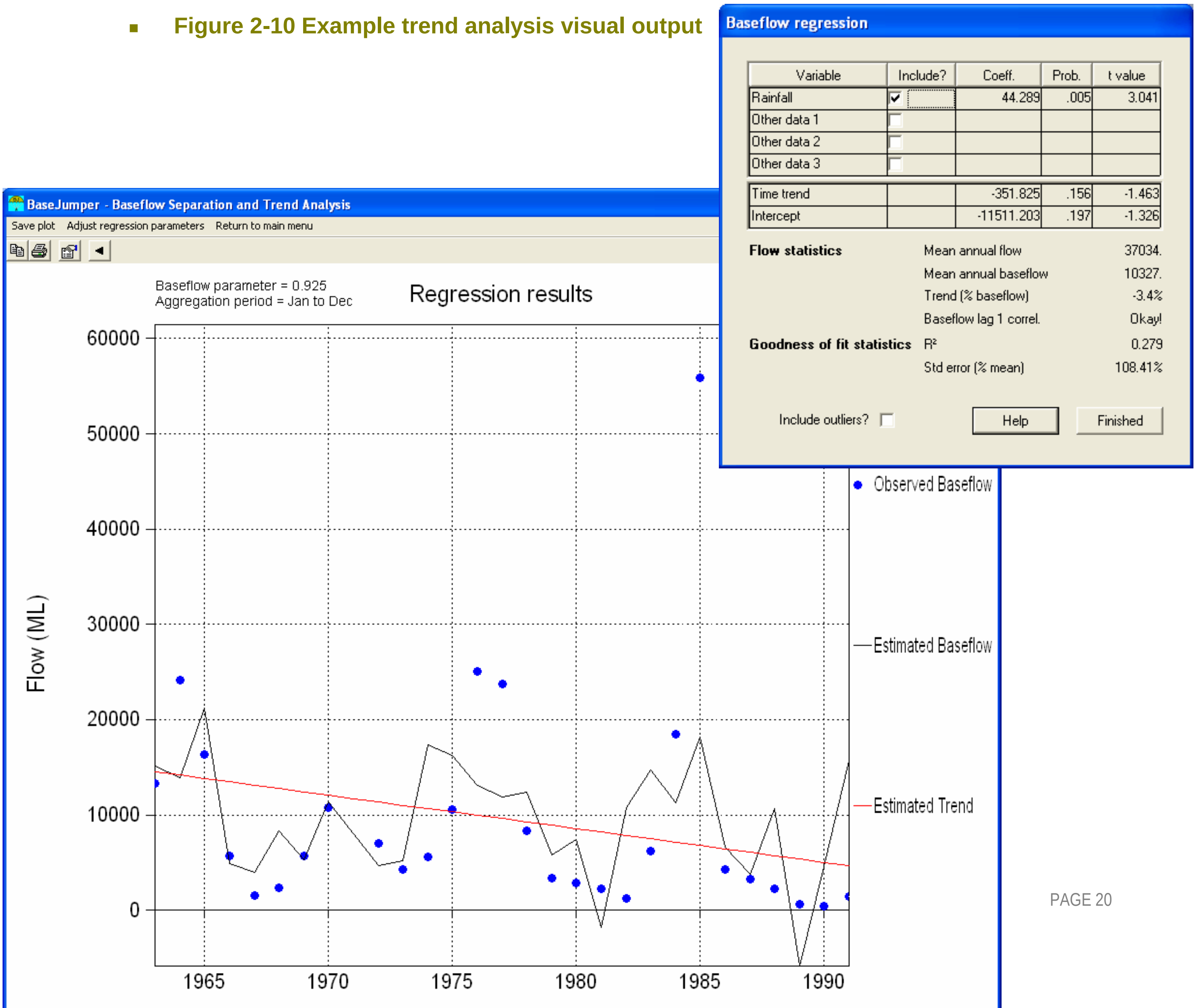
Equation 2-3

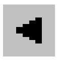
Where a_0 , a_1 and a_2 are coefficients.

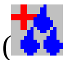
A visual output of the regression analysis shows the observed data (baseflow obtained through the Baseflow separation function of BaseJumper), the model fit, and the trend line (Figure 2-10). The regression parameters are also shown in a separate dialogue box.

The user should review this output to examine the trend in baseflow independent of climate before considering other possible influences on baseflow. The user is referred to Section 3 for a full explanation of the trend analysis results. In some cases, climate may explain all of the changes in baseflow that have occurred over time, and further analysis is not necessary.

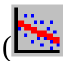
Figure 2-10 Example trend analysis visual output



Once the regression analysis has been reviewed, the Finished button on the parameter box can be selected followed by the Return to main menu option (or  icon) on the regression plot. Again, the user will be asked whether they would like to save the work. Relevant data, including input variable time series are saved, as well as statistics relating to the trend analysis, including R^2, regression coefficients, mean annual flow, mean annual baseflow, and the trend as a percentage of the mean annual baseflow. If required, the outputs of this text file can be transferred to Excel for further data analysis.

The user can input additional variables to be used in the trend analysis, if required. These additional data sets could include anthropogenic influences such as volumes of groundwater extracted and measurements of land use change. These additional variables must be input on an annual time step, which may require the user to aggregate data external to the BaseJumper program. It is necessary to aggregate these data sets on the same annual period as specified in the BaseJumper Analysis>Select Aggregation Period dialogue box. Up to three additional variables can be added, although it is not necessary to have this many variables to undertake the trend analysis. The Analysis>New additional variables menu option ( icon) is selected to input this additional data. For each additional variable, the user inputs the data using the same copy-and-paste process described for the input of streamflow and rainfall data (refer to Section 2.3). The user can specify the name of the input data series' by updating the "Other Data" headings in the baseflow regression dialogue box.

Once all additional variables have been added, the user selects Ok to return to the main menu.

Analysis>Baseflow regression with additional data ( icon) is again selected to analyse the impacts on baseflow trends resulting from the additional time series data added to the tool. A visual output of the regression analysis shows the original data, the model fit, and the trend line. The regression parameters are also shown in a separate dialogue box (Figure 2-11). This dialogue box allows the user to select which of the additional data sets should be included in the trend analysis. Only series' containing data will be available for selection. The trend analysis plot and statistics will be updated automatically when the series' for inclusion are modified.

The regression equation relevant to each analysis will depend on the independent variables included, however it will be of the form:

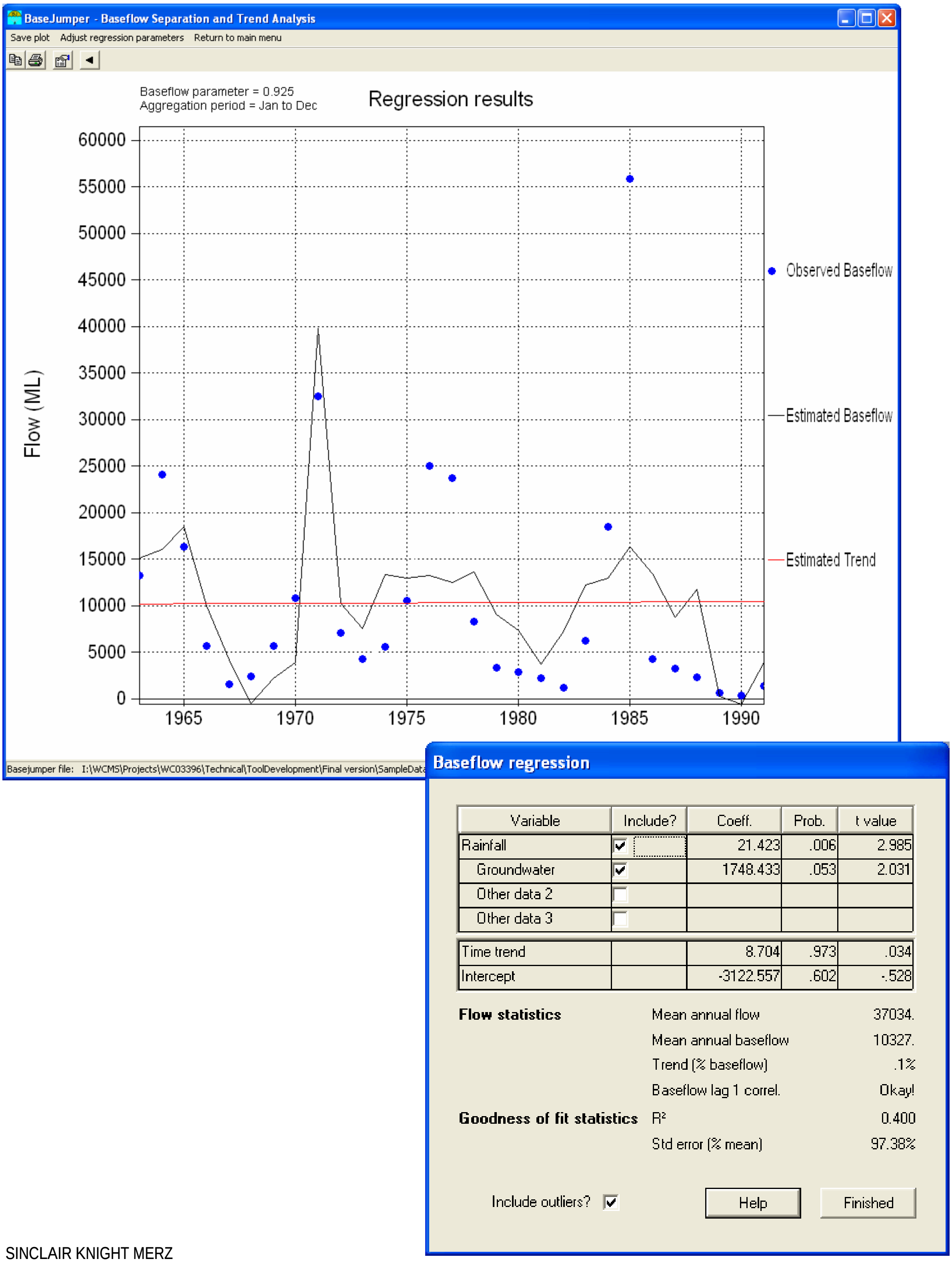
$$Baseflow = a_0 + a_1[Rainfall] + a_2[time] + a_3[OtherData1] + a_4[OtherData2] + a_5[OtherData3]$$

- **Equation 2-4**

Where a_0 , a_1 a_2, a_3, a_4 and a_4 are coefficients.

SKM

■ **Figure 2-11 Example trend analysis visual output**



Once the regression analysis has been finalised, the user selects Finished and Return to main menu (icon). The final regression analysis can be saved at this point. The results from the final trend analysis will be added to the output file.

BaseJumper automatically retains the information from the last regression, should the user want to return to this analysis.

2.8 BaseJumper Help Options

The Help menu provides the user with some additional information about the BaseJumper program.

Help>Help (or the icon) opens the help contents file, which contains information from the user manual as relevant to the function of the program. Information on specific aspects of the program can also be accessed while using the software, by selecting the Help button on the various dialogue boxes.

Help>About provides information about the BaseJumper development.

Help>Browse Text File (icon) allows the user to open a text file to view and edit. This text editor has all the functionalities of NotePad.

3. Interpreting BaseJumper Results

BaseJumper provides both visual and text outputs for interpretation. These should be examined thoroughly to understand the results of the analysis. Care should be taken in the analysis to ensure that the results are interpreted in a meaningful way.

During the baseflow separation stage, outputs are presented visually on the screen, which provides an opportunity for the user to confirm the resulting timeseries before proceeding. Section 2.5.2 outlined the key features to consider for reasonable baseflow estimation. The daily streamflow and resulting baseflow data is provided in the text file output. Additionally, the digital recursive filter parameter applied is noted, and the resulting Baseflow Index (BFI) is also provided. The BFI is calculated as the total baseflow divided by the total streamflow from the entire dataset.

The regression analysis results are initially provided for a regression analysis using only rainfall as the independent variable. Further analysis may be undertaken with additional independent variables, and outputs for these additional analyses are subsequently provided.

The visual outputs provide basic information on the regression analysis, primarily the change in baseflow over time and the trend associated with this change. A good model fit occurs when the fitted model (black line in the BaseJumper regression analysis plots) can replicate the observed data (blue dots). Both high and low flow events as well as fluctuations in the data should be reproduced. For regression models where the fitted and the observed data match well, the trend in the dependent variable (baseflow) can be assessed. The model trend line indicates the general direction of the trend. A positively sloping line implies there has been an increase in baseflow over time, while a negatively sloping line infers declining baseflow over time. A relatively flat line suggests that there is has been no change in baseflow contributions to streamflow, and that the independent variables (rainfall, plus any additional variables added by the user) are able to fully account for any fluctuations in baseflow at this location. All of these trend results are equally valid and provide useful information about the interaction between local groundwater and surface water systems. However, it is still necessary for the user to fully consider BaseJumper results in light of the catchment characteristics to ensure the results are feasible. To assist in this, additional information about the regression analysis is provided in the form of statistical outputs. The analysis of these statistical outputs is discussed below.

Statistical outputs relating to the regression analysis are shown in a dialogue box which is available at the same time as the visual results are displayed. The coefficients relating to each variable are shown, along with the t-statistic and probability. The coefficients in this dialogue box correspond with the values represented by a_n in Equation 2-3 and Equation 2-4. The statistical significance of each term in these equations is shown by the t value and probability terms. The t value is computed by dividing the estimated value of the parameter by its standard error. If this value is sufficiently different from zero, the test is considered to be statistically significant. The Prob(t) value determines the level of statistical significance. The smaller the value of Prob(t), the more significant the parameter. For example, a Prob(t) value of 0.05 indicates that there is only a 0.05 (5%) chance that the estimated value of the parameter value could be obtained by chance alone. This represents statistical significance at the 5% level, which relates to a strongly evident trend. A Prob(t) value of 0.1 represents statistical significance at the 10% level, which reflects a moderately evident trend. Values of Prob(t) greater than 0.1 indicates that any trend identified can not be attributed to an actual trend and may be due to random variability in the data.

Goodness of fit and flow statistics are also shown on the dialogue box. These statistics summarise the:

- R^2: this denotes the percentage variation in the dependent variable (baseflow) accounted for by the independent variables (rainfall, plus any other additional variables). Values closer to 1 represent a better model fit.
- Coefficient of efficiency: used to assess the predictive power of hydrological models. This is calculated as $E = 1 - \frac{\sum_{t=1}^{T}(Q_o^t - Q_m^t)^2}{\sum_{t=1}^{T}(Q_o^t - \overline{Q}_o)^2}$ where Q_o is observed discharge, and Q_m is modelled discharge. Q^t is discharge at time *t*. A value of zero for the coefficient of efficiency indicates that the observed mean is as good a predictor as the model, while negative values indicate that the observed mean is a better predictor than the model. Values closer to 1 represent a better model fit.
- Standard error: standard deviation of a distribution of sample means. The lower the standard error, the better the model fit.
- MAF: mean annual flow (ML)
- MABF: mean annual baseflow (ML)
- Trend/MABF: annual trend as a proportion of the mean annual baseflow (%/year)

The text output saved by the user also summarises the information obtained through the application of BaseJumper. Figure 3-1 shows an example text output from BaseJumper. The text output records the date of the save, and documents the daily datasets used in the baseflow separation, as well as the annual datasets used in the regression analysis (if undertaken). The statistical outputs associated with the regression analysis are also provided. The following information is provided:

- Date saved
- BaseJumper version

- Regression results
  - Variables used in the regression analysis
  - Coefficients, t-statistic and Prob(t) associated with each variable applied
  - Whether outliers have been included in the analysis or not
  - Number of years of data
  - Nsample: number of data points used in the analysis. Years with significant missing data will be removed from analysis, making Nsample different to the number of years of data.
  - NumUseVar: Number of variables used in the analysis
  - AvEst: Average of the estimated data
  - AvObs: Average of the observed data
  - Ssreg: sum of squares for the regression
  - Ssresid: sum of squares for the residuals
  - R^2
  - SE: Standard error
  - CE: Coefficient of efficiency
  - Mean annual flow
  - Mean annual baseflow
  - Baseflow trend
  - Trend as a proportion of mean annual baseflow
  - Probability that the trend is not significant
- Baseflow separation data
  - Digital filter parameter applied
  - Number of passes
  - Baseflow index (BFI)
- Annual data
  - Years of data in record
  - Start and end months for the aggregation of data
  - Number of months in each aggregation period
  - Annual timeseries of flow, baseflow, rainfall, and any other additional data added by the user for the regression analysis
- Daily data
  - Days of data in record

- Daily timseries of flow, baseflow and rainfall. The date column associated with the daily data is presented in the YYYYMMDD format, where YYYY represents the year, MM represents the numerical value of the month (01-12) and DD represents the day of the month.

SKM

- **Figure 3-1 Example BaseJumper data and results file (note that some annual and daily timeseries data has been removed for practicaility)**

```
================================
BaseJumper data and results file
================================

Date saved:        16 February 2007 12:15
Program version:   Version 0.4, 14th February
Analysis status:   4

Regression results
------------------

        Variable Y/N    Coefficient   tstat    tprob
    Other data 1 1            58.81  68.408    0.000
    Other data 2 0          1439.07   0.000    0.000
    Other data 3 0             0.00   0.000    0.000
           Trend 1          1451.10   1.832    0.104
     Coefficient 1        -29578.24  -1.096    0.305

        numyears                              24
         nsample                              23
       NumUseVar                               2
           avest                       68390.102
           avobs                       68390.109
           ssreg                 27571173400.000
         ssresid                  5725342720.000
             sst                           0.000
              R2                           0.828
              SE                       17358.967
              ce                           0.172

       Mean Annual Flow (MAF):         111499.6  ML/yr
  Mean Annual Baseflow (MABF):          66575.7  ML/yr
               Baseflow trend:            350.4  ML/yr
        Baseflow trend / MABF:             52.6%
Prob trend is not significant:             56.3%


Baseflow Separation
--------------------

Digital filter parameter:0.925
        Number of passes:    3
    Baseflow Index (BFI): 0.59


Annual Data
-----------

    years of data: 24

  Year          Flows        Baseflows    Other data 1        Lin.trend
  1980       70017.99         46140.89         -999.00                1
  1981      174787.23        115016.94            3.89                2
  1982      115951.20         75753.91            0.44                3
  1983       66217.28         42711.53            1.37                4

Etc…

Daily Data
----------

     days of data: 9002

YYYYMMDD           Flows        Baseflows            Rain
19800101           81.64            57.89            1.40
19800102           70.55            59.05            1.20
19800103           71.96            59.97            0.40
19800104           78.57            61.02           12.00

Etc…
```

4. Advanced application of BaseJumper

The application of BaseJumper described in the previous sections is suitable for most purposes. In particular, it covers the use of the tool in cases where the user has streamflow data at an unregulated site (free of regulating structures, diversions and discharges). However, in many locations it may not be possible to obtain information for a site that is unaffected by regulating structures, diversions and discharges. In some cases, this may not prohibit the application of BaseJumper. This section describes the data processing steps required to enable users to apply the tool to river reaches affected by regulating structures, diversions and discharges.

It should be noted that the advanced application of BaseJumper is recommended only for users with hydrological and hydrgeological knowledge of the site being studied. The methods described in this section assume a suitable level of understanding of the physical processes occurring at the site.

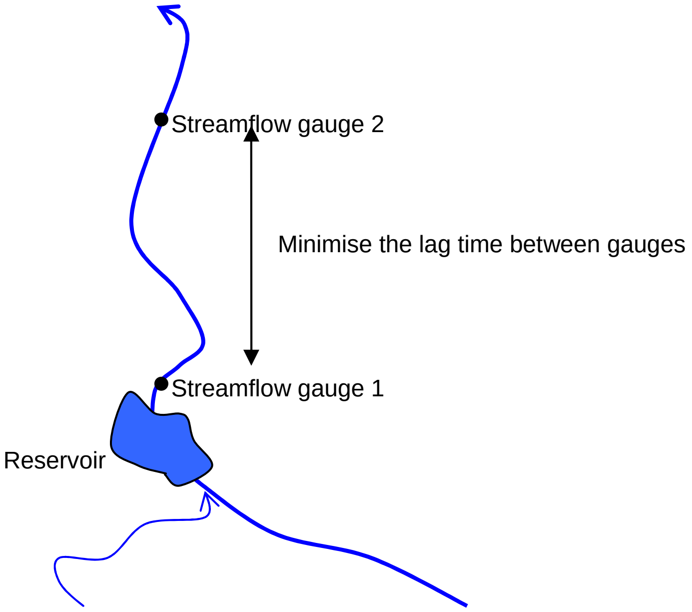
4.1 Adjustment of data to account for regulating structures

Regulating structures such as dams and reservoirs impact on streamflows downstream of the structure by controlling the volume of water released. If this flow regulation is not taken into account, the streamflow time series is not able to be used for baseflow separation as described in Section 2. In order to generate a sensible baseflow time series, this flow regulation must be considered.

Streamflow regulation can be taken into account in situations where two streamflow gauges are available on the reach of interest. Both gauges must be downstream of the regulating structure, with the second gauge being further downstream than the first. Ideally, the gauges should be located less than 1 day travel time (streamflow travel time) of each other. Longer travel times are acceptable but may have a reduced chance of success, as the lag between the gauges must be accounted for. Figure 4-1 displays a schematic diagram of streamflow gauge locations for the analysis of a regulated river reach.

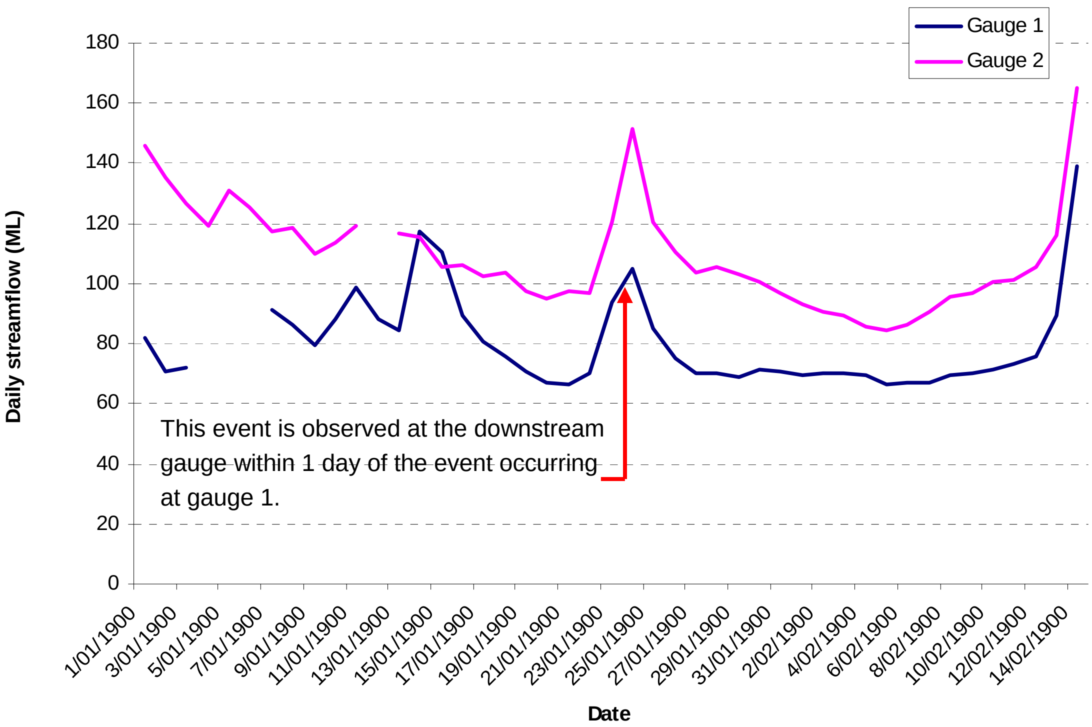
SKM

- **Figure 4-1 Schematic diagram of streamflow gauge location for analysis of regulated river reach**

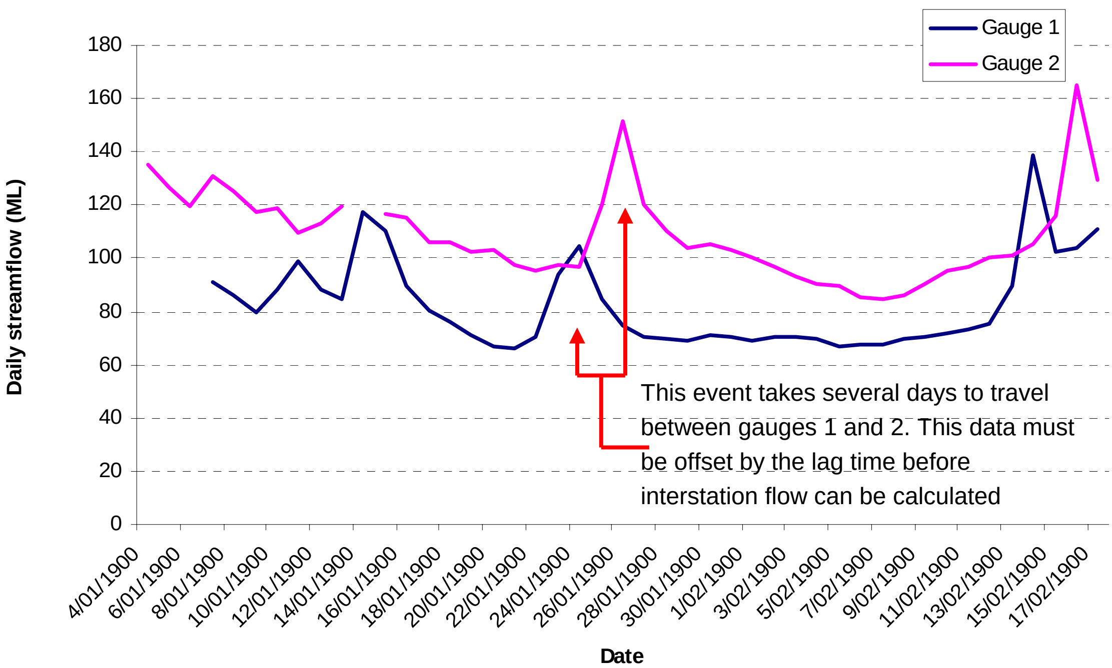


In order to determine the lag time between the two streamflow gauges, the hydrographs should be compared. To do this, the daily time series data should be plotted in Excel or some other similar data management software. For the lag time between the two gauges to be considered less than one day, streamflow events should be observed at gauge 2 within one day of the event occurring at gauge 1. Where the travel time between the two gauges is greater than 1 day, this lag must be accounted for. One data set should be transposed by the appropriate number of days before the interstation flow is calculated. Figure 4-2 and Figure 4-3 provide example streamflow hydrographs for two sample gauges which are positioned as shown in schematic of Figure 4-1. These figures are annotated to indicate how the hydrographs can be used to determine the travel time between streamflow gauges.

- **Figure 4-2 Daily streamflow hydrographs for streamflow gauges within 1 day travel time**



- **Figure 4-3 Daily streamflow hydrographs for streamflow gauges with travel time of a day or more**



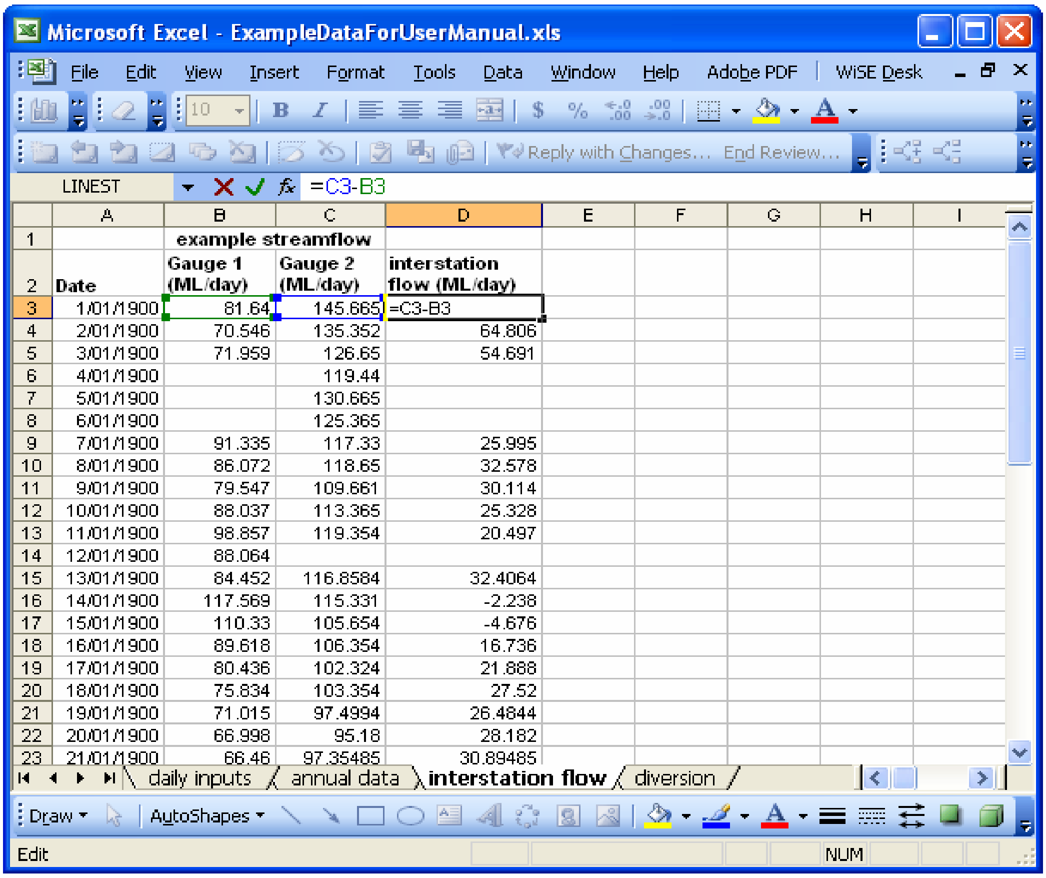
The interstation flow must be calculated using the two streamflow gauges. Baseflow separation can then be undertaken using the interstation flow, to determine the baseflow entering the stream in the reach between the two gauges. A daily time series of interstation flow is calculated by subtracting the daily streamflow at gauge 1 from the daily streamflow at gauge 2. That is:

Interstation flow $= Q_2 - Q_1$

Where Q_1 is the daily streamflow at gauge 1 (downstream of the regulating structure) and Q_2 reflects the daily streamflow at gauge 2 (located downstream of both the regulating structure and streamflow gauge 1).

Where the lag time between the two gauges is less than one day, there is no need to take the streamflow travel time into account. Consequently, the interstation flow on any given day is directly calculated by considering the streamflow at each gauge on the same day. Missing data in the daily time series data for either gauge 1 or 2 should be reflected in the interstation flow time series. Figure 4-4 provides an example of this calculation using sample data.

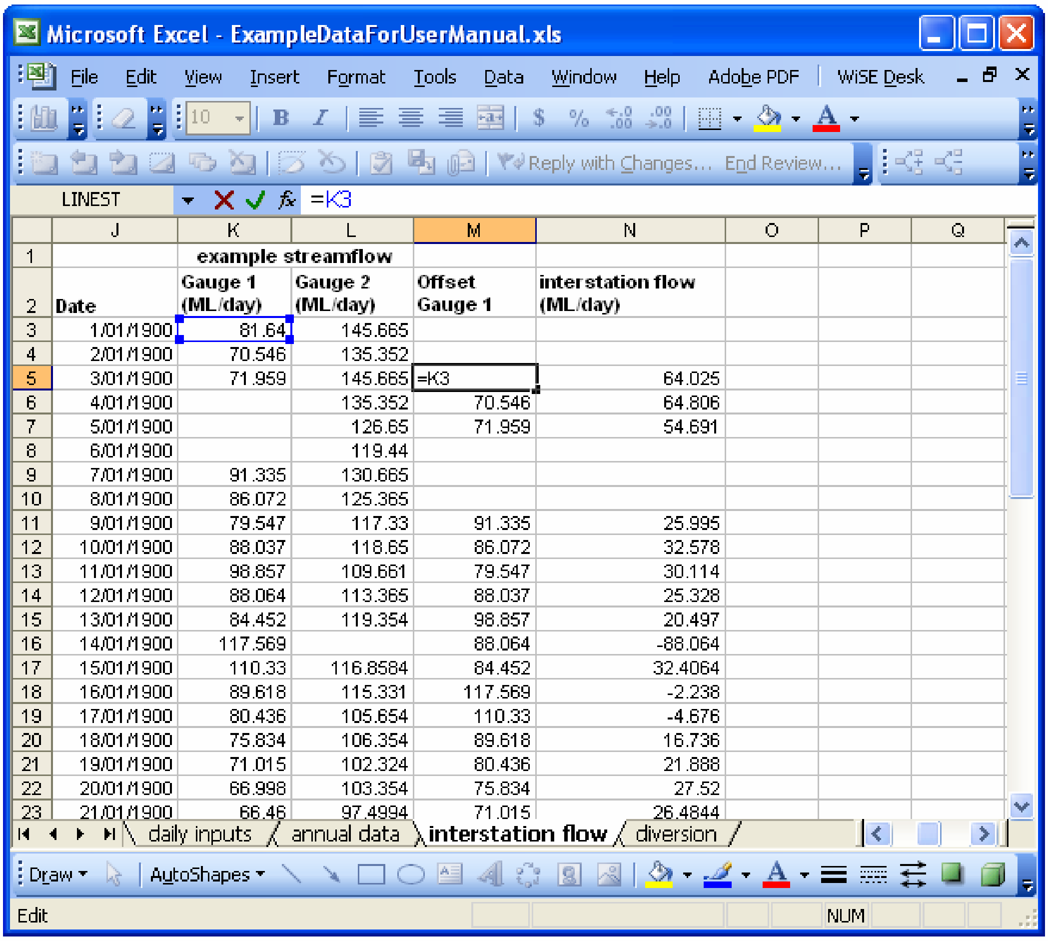
- **Figure 4-4 Calculation of interstation flow (travel time between gauges < 1day)**



	A	B	C	D
1		example streamflow		
2	Date	Gauge 1 (ML/day)	Gauge 2 (ML/day)	interstation flow (ML/day)
3	1/01/1900	81.64	145.665	=C3-B3
4	2/01/1900	70.546	135.352	64.806
5	3/01/1900	71.959	126.65	54.691
6	4/01/1900		119.44	
7	5/01/1900		130.665	
8	6/01/1900		125.365	
9	7/01/1900	91.335	117.33	25.995
10	8/01/1900	86.072	118.65	32.578
11	9/01/1900	79.547	109.661	30.114
12	10/01/1900	88.037	113.365	25.328
13	11/01/1900	98.857	119.354	20.497
14	12/01/1900	88.064		
15	13/01/1900	84.452	116.8584	32.4064
16	14/01/1900	117.569	115.331	-2.238
17	15/01/1900	110.33	105.654	-4.676
18	16/01/1900	89.618	106.354	16.736
19	17/01/1900	80.436	102.324	21.888
20	18/01/1900	75.834	103.354	27.52
21	19/01/1900	71.015	97.4994	26.4844
22	20/01/1900	66.998	95.18	28.182
23	21/01/1900	66.46	97.35485	30.89485

Where the travel time is 1 day or more, the data must be offset by the travel time before the interstation flow can be calculated. To do this, shift the data for gauge 1 by the number of days travel time between gauges, so that the flow events are occurring at the same time. For example, using the data from Figure 4-3, gauge 1 data has been shifted by 2 days (see Figure 4-5). Interstation flow can be calculated as above.

- **Figure 4-5 Calculation of interstation flow (travel time between gauges > 1day)**



	J	K	L	M	N
1		example streamflow			
2	Date	Gauge 1 (ML/day)	Gauge 2 (ML/day)	Offset Gauge 1	interstation flow (ML/day)
3	1/01/1900	81.64	145.665		
4	2/01/1900	70.546	135.352		
5	3/01/1900	71.959	145.665	=K3	64.025
6	4/01/1900		135.352	70.546	64.806
7	5/01/1900		126.65	71.959	54.691
8	6/01/1900		119.44		
9	7/01/1900	91.335	130.665		
10	8/01/1900	86.072	125.365		
11	9/01/1900	79.547	117.33	91.335	25.995
12	10/01/1900	88.037	118.65	86.072	32.578
13	11/01/1900	98.857	109.661	79.547	30.114
14	12/01/1900	88.064	113.365	88.037	25.328
15	13/01/1900	84.452	119.354	98.857	20.497
16	14/01/1900	117.569		88.064	-88.064
17	15/01/1900	110.33	116.8584	84.452	32.4064
18	16/01/1900	89.618	115.331	117.569	-2.238
19	17/01/1900	80.436	105.654	110.33	-4.676
20	18/01/1900	75.834	106.354	89.618	16.736
21	19/01/1900	71.015	102.324	80.436	21.888
22	20/01/1900	66.998	103.354	75.834	27.52
23	21/01/1900	66.46	97.4994	71.015	26.4844

Typically, the streamflow at gauge 2 will be greater than that at gauge 1. However, there may be instances where this is not the case. In some cases, the interstation flow may be infrequently negative. In these situations, the interstation flow can still be used in BaseJumper and periods of 'negative flow' will be treated as valid data. However, baseflow will not be calculated for these days. For reaches where the river is considered to be a losing stream it would be likely to expect long periods of negative interstation flows. In such cases, baseflow would not exist (baseflow is defined as the component of streamflow that is sourced from groundwater aquifers - by definition, it can not flow in the opposite direction), and an alternative site should be considered.

Periods of missing data in either gauge timeseries should be reflected in the interstation flow.

The calculated interstation flow time series can be used directly as input into BaseJumper for baseflow separation. The steps outlined in Section 2 can be followed using the interstation flow for analysis. Results obtained using this data will reflect the trends in baseflow in the river reach between gauge 1 and 2, and the influence of the regulating structure will have been removed.

It should be noted that flow monitoring errors are likely to compound when using this approach and hence there is much scope for variability and uncertainty in baseflow estimates. Theoretically, baseflow should be relatively stable (or follow an exponential decay function in the absence of catchment rainfall) on successive days, so comparison of baseflow estimates on successive days will give an indication of the uncertainty in those estimates.

4.2 Magnifying trends in unregulated river reaches

Trends in unregulated reaches may be magnified by considering the interstation flow between two streamflow gauges. This eliminates the ‘noise’ in the data by considering only the baseflow generated within the interstation reach. Interstation flows should be calculated and applied as described in Section 4.1.

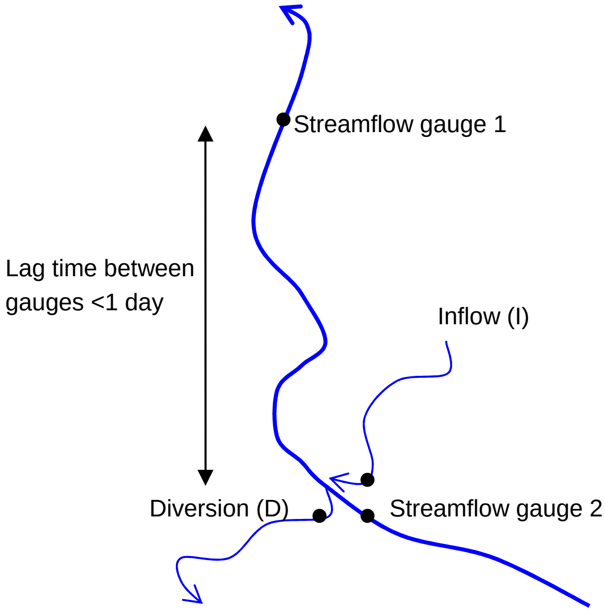
4.3 Adjustment of data to account for diversions and inflows

Diversions, discharges and inflows impact on streamflow by removing or adding water to the river. This modification of the natural flows must be taken into account before baseflows can be separated from the streamflow data.

Figure 4-6 shows a schematic diagram of the streamflow gauge information required to analyse a river reach influenced by diversions or inflows. Two streamflow gauges are required on the main stream. The diversion and/or inflow data is also required.

SKM

Figure 4-6 Schematic diagram of streamflow gauge location for analysis of streams with diversions



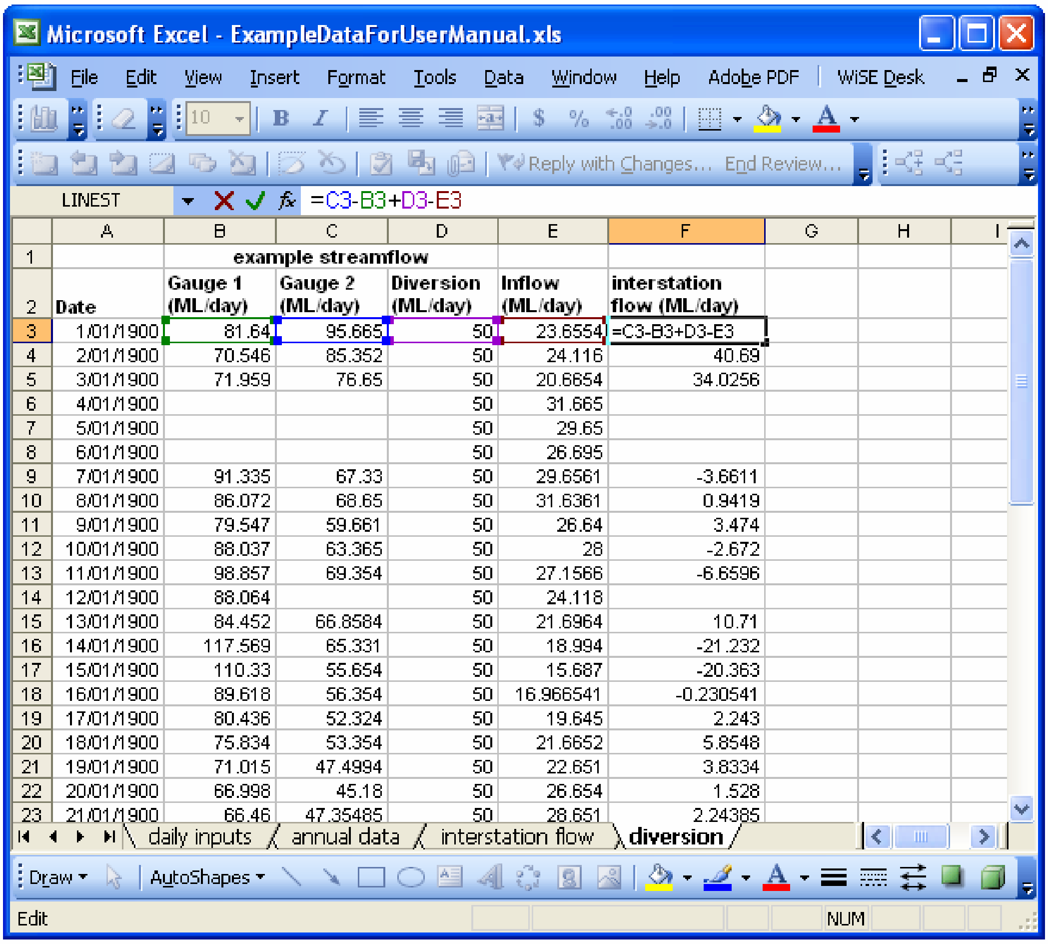
The main stream data must be adjusted to take the diversion/inflow into account. To do this, the interstation flow must be calculated, and the diversions/inflows accounted for. That is:

Interstation flow $= Q_1 - Q_2 + D - I$

Where Q is the daily streamflow on the main reach, D is the diversion volume and I is the inflow or discharge volume. Using this method, it is assumed that all gauging sites are located within 1 day travel time of each other. If there is a greater lag time, this must be considered as described in Section 4.1. Missing data in any of these inputs should be reflected in the final interstation flow time series. Figure 4-7 provides an example of the calculation for a 50 ML/day diversion and a small inflowing creek.

SKM

- **Figure 4-7 Adjustment of data to account for diversions and inflows**



	A	B	C	D	E	F
1		example streamflow				
2	Date	Gauge 1 (ML/day)	Gauge 2 (ML/day)	Diversion (ML/day)	Inflow (ML/day)	interstation flow (ML/day)
3	1/01/1900	81.64	95.665	50	23.6554	=C3-B3+D3-E3
4	2/01/1900	70.546	85.352	50	24.116	40.69
5	3/01/1900	71.959	76.65	50	20.6654	34.0256
6	4/01/1900			50	31.665	
7	5/01/1900			50	29.65	
8	6/01/1900			50	26.695	
9	7/01/1900	91.335	67.33	50	29.6561	-3.6611
10	8/01/1900	86.072	68.65	50	31.6361	0.9419
11	9/01/1900	79.547	59.661	50	26.64	3.474
12	10/01/1900	88.037	63.365	50	28	-2.672
13	11/01/1900	98.857	69.354	50	27.1566	-6.6596
14	12/01/1900	88.064		50	24.118	
15	13/01/1900	84.452	66.8584	50	21.6964	10.71
16	14/01/1900	117.569	65.331	50	18.994	-21.232
17	15/01/1900	110.33	55.654	50	15.687	-20.363
18	16/01/1900	89.618	56.354	50	16.966541	-0.230541
19	17/01/1900	80.436	52.324	50	19.645	2.243
20	18/01/1900	75.834	53.354	50	21.6652	5.8548
21	19/01/1900	71.015	47.4994	50	22.651	3.8334
22	20/01/1900	66.998	45.18	50	26.654	1.528
23	21/01/1900	66.46	47.35485	50	28.651	2.24385

Typically, this calculation to estimate the interstation flow will result in positive values. However, there may be instances where this is not the case. In some cases, the interstation flow may be infrequently negative. In these situations, the interstation flow can still be used in BaseJumper and periods of ‘negative flow’ will be treated as valid data, but baseflow will not be calculated for these days. However, for reaches where the river is considered to be a losing stream it would be likely to expect long periods of negative interstation flows. In such cases, baseflow would not exist (baseflow is defined as the component of streamflow that is sourced from groundwater aquifers - by definition, it can not flow in the opposite direction), and further analysis of the data is not necessary as baseflow separation and trend analysis can not be undertaken.

The interstation flow can be used in BaseJumper. It should be directly provided as the input streamflow time series, as described in Section 2.3. The subsequent approach to obtain information on baseflow trends is then as described in the remainder of Section 2. Results obtained using this data will reflect the trends in baseflow in the river reach between the two main stream gauges.

It should be noted that flow monitoring errors are likely to compound when using this approach and hence there is much scope for variability and uncertainty in baseflow estimates. Theoretically, baseflow should be relatively stable (or follow an exponential decay function in the absence of catchment rainfall) on successive days, so comparison of baseflow estimates on successive days will give an indication of the uncertainty in those estimates.

5. References and Further Reading

Cook, R.D. (1977) *Detection of Influential Observation in Linear Regression.* Technometrics, Vol. 19, No. 1, pp. 15-18. February 1977.

Cook, R.D. (1979) *Influential Observations in Linear Regression*. Journal of the American Statistical Association, Vol. 74, No. 365. (Mar., 1979), pp. 169-174. March 1979.

Daamen, C., McAuley, C. and Durkin, P. (2006) *Effects of urban development on groundwater baseflow in streams*. Proceedings of Enviro06 Conference, 9-11 May 2006, Melbourne, Australia.

Nathan, R.J. and McMahon, T.A. (1990) *Evaluation of Automated Techniques for Base Flow and Recession Analyses*. Water Resources Research Vol 26, Number 7, pages 1465-1473. July 1990.

Nathan, R.J., Nandakumar, N., and Smith, W.E. (1999) *On the Application of Generalised Additive Models to the Detection of Trends in Hydrologic Time Series Data*. Paper submitted to the Hydology and Water Resources Symposium, Brisbane, July 1999.

Neal, B.P., Nathan, R.J. and Evans, R. (2004) *Survey of Baseflows in Unregulated Streams of the Murray-Darling Basin*. 9th Murray-Darling Basin Groundwater Workshop, 17-19 Feb 2004, Bendigo, Victoria.