USER MANUAL FOR

WHAT

Well Hydrograph Analysis Toolbox

July 23, 2015

Document up to date for software version 4.1.7-beta

Jean-Sébastien Gosselin[1], Christine Rivard[2], and Richard Martel[1]

[1] Institut national de la recherche scientifique, Centre Eau Terre Environnement, 490 rue de la Couronne, Quebec City, Quebec, Canada

[2] Geological Survey of Canada, Quebec Division, 490 rue de la Couronne, Quebec City, Quebec, Canada

License

Contents

List of Figures

1 Introduction

1.1 What is WHAT

WHAT (Well Hydrograph Analysis Toolbox) is a free, open source, and cross-platform interactive computer program whose main focus is the interpretation of observation well hydrographs. It is written in the Python 2.7 programming language and is currently maintained and developed by Jean-Sébastien Gosselin at INRS-ETE (`www.ete.inrs.ca`). The source code and a stand-alone executable for Windows 7 are available free of charge for download on GitHub (`www.github.com/jnsebgosselin/WHAT`).

If you encounter any problems or errors during program execution, have any questions, or have suggestions on how to improve WHAT, please contact Jean-Sébastien Gosselin at this email address: jnsebgosselin@gmail.com.

1.2 Features

Below are listed the features that are available in the current version of WHAT, as well as the ones that are going to be incorporated in future versions of the software.

Features available in current version of WHAT :

- Preparation of Gapless Daily Weather Time-series :
  - Graphical interface to the online Canadian Daily Climate Database (CDCD) to search for weather station by location coordinate.
  - Automatic downloading and formatting of available data from the CDCD.
  - Estimation of missing data and automatic gapfilling of the datasets.
- Data Visualization:
  - Data exploration in a user-friendly and dynamic graphical environment.
  - Production of publication-quality graphs in vectorial format.
- Weather Data Analysis :
  - Estimation of yearly and monthly normals.
  - Estimation of the daily potential evapotranspiration.
- Water Level Analysis :

  – Calculation of the master recession curve (MRC).
  – Estimation of groundwater recharge from the MRC using a continuous water-table fluctuation (WTF) method.

Features planned for future versions of WHAT :

- Synthetic hydrographs production for the estimation of groundwater recharge and prediction of water levels.
- Assessment of the level of confinement of the aquifer at the well location based on a barometric response function analysis of the water level record.

1.3 Installation

WHAT can run on Windows, Linux, or OS X computer operating systems. However, a stand-alone executable of the program is currently released and tested only for the Windows 7 platform. This executable should also be compatible with Windows XP. For the Linux and OS X platforms, the software can be run directly from the source code, provided that Python 2.7 and all the required third party packages are installed on the computer (PySide, NumPy, matplotlib, xlrd, xlwt).

The stand-alone executable for Windows 7 is distributed in a Zip archive that can be downloaded freely on GitHub (https://github.com/jnsebgosselin/WHAT/releases). This archive contains:

- the GNU General Public License;
- a folder named "WHAT" that contains all the necessary system files for the program to run, including the file "WHAT.exe" from which the software can be started;
- a folder named "Projects" where all input and output files used or created by WHAT are stored by default. This folder includes samples of input and output files that provide a quick and convenient way to test and learn the various features of the program.

Once the content of the Zip archive has been extracted, the program can be started directly from the WHAT.exe executable file that is contained within the folder named WHAT. The software can conveniently run from any location on the computer or from any storage device without the need to install the program beforehand.

1.4 Overview of the Graphical User Interface

The Graphical User Interface (GUI) of WHAT mainly consists of a menu bar, a console area, and a central view panel (Fig. 1.1). The *menu bar* is located in the top right corner of the GUI. This is where you can view the name of the current project, open an already existing project or create a new one. The *console* is located at the bottom of the GUI and is used to report technical information about the various tasks accomplished by the program as well as warning and error messages. The console can be collapsed to save space, or can be extended to the entire window area. The *central view panel* is the main component of the GUI. It is where the various features of the software are displayed. The content of this panel is divided into four tabs: *Download Data*, *Fill Data*, *Hydrograph*, and *About*. These tabs are described in more details below and are shown in Fig. 1.2.

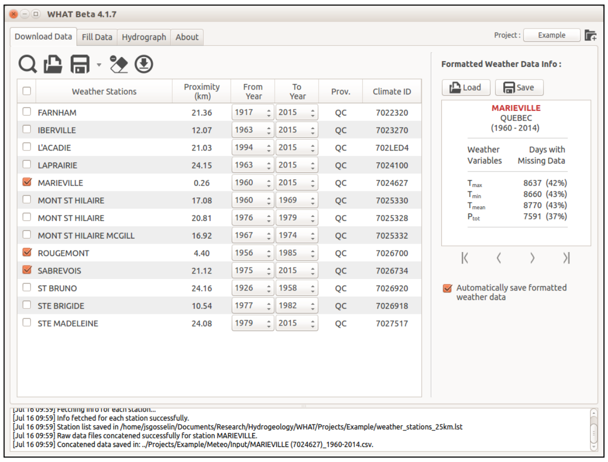
Download Data : This tab (Fig. 1.2a) provides a graphical interface to the online Canadian Daily Climate Database (CDCD), owned and operated by Environment Canada, from which it is possible to interactively search for stations by location coordinates, download the available data, and automatically organize the data in a format compatible with WHAT.

Fill Data : This tab (Fig. 1.2b) provides an automated procedure to estimate and fill the missing values in the daily weather datasets. Missing data from a given station are estimated with data from selected neighboring weather stations using a multiple linear regression model.

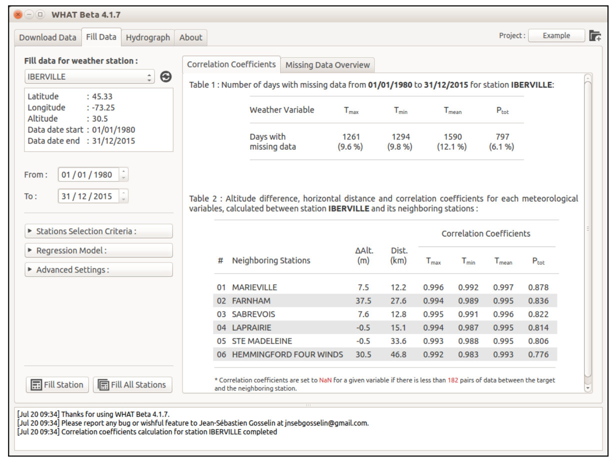
Hydrograph : This tab is used for viewing and plotting both groundwater level and weather time series. For this purpose, two modes are available: the *layout* and the *computation* mode. Both modes share the same weather and water level datasets and it is possible to switch from one mode to the other at anytime. The **layout** mode (Fig. 1.2c) provides a graphical interface to interactively produce publication-quality graphs. The **computation** mode (Fig. 1.2d) consists in a dynamic graphical environment where data can be visualized, manipulated and analyzed. Various computational tools are available in this mode, including the estimation of the hydrograph Master Recession Curve (MRC) and the estimation of groundwater recharge.

About : This tab (Fig. 1.1) displays copyright, licensing and general information about WHAT.

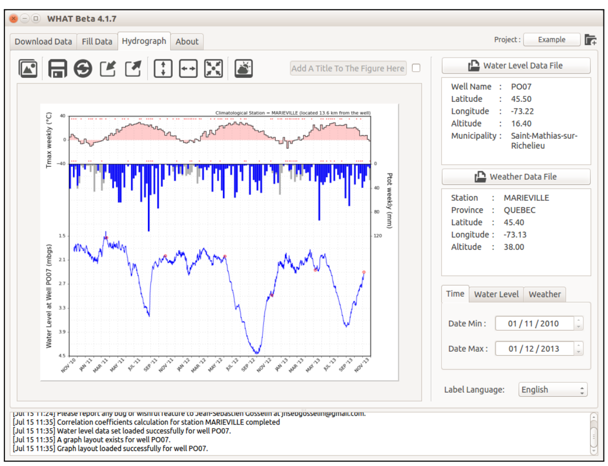
Figure 1.1: Screenshot of WHAT v4.1.7-beta in Ubuntu-Linux 15.04 showing the *About* tab.



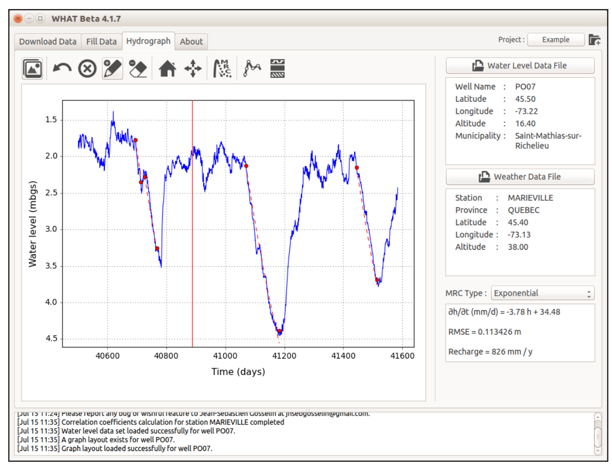
(a) *Download Data* tab.



(b) *Fill Data* tab.



(c) *Hydrograph* tab in *Layout* mode.



(d) *Hydrograph* tab in *Computation* mode.

Figure 1.2: Screenshots of WHAT v4.1.7-beta captured in Ubuntu Linux 15.04. showing : (a) the *Download Data* tab, (b) the *Fill Data* tab (c) the *Hydrograph* tab in *Layout* mode, and (d) the *Hydrograph* tab in *Computation* mode.

2 Data Management by Projects

2.1 Introduction

Data is managed in WHAT by project. That is all input and output files relative to a given project are saved within a common folder called the "project folder". This file management system allows you to easily backup and move your projects from one location to the other since all the files relating to a given project are saved at the same place.

On first launch, WHAT will automatically open an example that is distributed with the software with all the necessary files to easily and quickly test the functionality of the program. The title of the currently opened project is shown in the menu bar at the top of the interface. Additional information about the project can be displayed by clicking on the small "i" icon located next to the project name. There can be only one opened project at a time per instance of WHAT.

2.2 Create a New Project

To start a new project, click on the button *New Project...* with the small folder icon located at the right end of the WHAT menu bar (see Figure 1.1). This will open a new dialog window (see Figure 2.1) where you can enter various information about your project such as its title, author and location coordinates.

Clicking on the button *Save* creates a new project folder named after your project title in which your project information are saved in a file with a ".what" extension. The new folder is created in the location defined by the *Save in Folder* directory path. For example, saving the *My New Project* of Figure 2.1 would create a folder named "My New Project" in the directory "*C:\Users\johndoe\WHAT_4.0.5-beta\Projects*" and would saved the project information in the file named "my_new_project.what". It is possible to change the directory where the project folder is created by clicking the small folder icon located next to the *Save in Folder* directory path.

2.3 Open a Project

To open a new project, click on the name of the currently opened project in the menu bar at the top of WHAT window. This will open a new dialog window where you can browse your folders to select an already existing project file (*.what), and then click Open. WHAT will then open the project and the currently opened project displayed in the menu bar should change to the name of the project you just selected.

The path to your project folder is stored in WHAT in a relative format. This means that if you change the location of your project folder relative the "WHAT.exe", your will have to redirect

WHAT to the new location of your project by repeating the procedure described in the paragraph above.

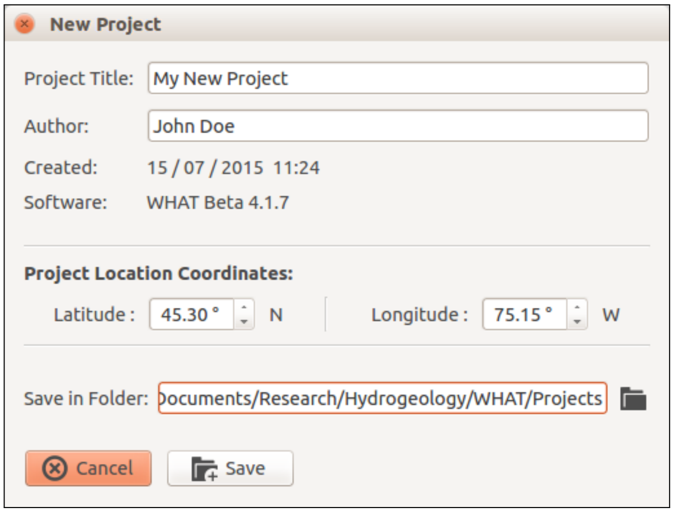


Figure 2.1: Screenshot of WHAT v4.1.7-beta in Ubuntu-Linux 15.04 showing the *New Project* dialog window.

2.4 Project Folder Structure Overview

In addition to the project file (.what file extension) that is created when saving a new project, WHAT automatically generates various files and sub-folders that are required for it to run. This file organization is briefly described here and an example is presented in Figure 2.2. The project folder contains two sub-folders named "Meteo" and "Waterlvl" in addition to a few other files.

Meteo The folder *Meteo* contains three sub-folders named respectively Raw, Input and Output. The folder **Raw** is where are saved the weather data files downloaded from the CDCD. These are coma-separated values (CSV) files that contain weather data on a yearly basis. All the data files for a given weather station are saved within a common folder named after the station name and its unique identification number (IDN). For example, in Figure 2.2, the raw data file "eng-daily-01011980-12311980.csv" that contains weather data of the station "Marieville" for the year 1980 is saved within a folder named "MARIEVILLE (7024627)" where the number in parentheses is the unique IDN of the station.

The folder **Input** contains the formatted weather data files produced from the raw data files. These are tab-separated values (TSV) files that are named after the station's name and IDN.

The folder **Output** is where are saved the gapless weather time-series produced from the content of the Input folder. These are saved in TSV text files with the extension ".out". The files with the extension ".log" are TSV text files that contain detailed information about every missing daily weather value that were estimated by the program to produce the gapless time-series (.out files).

Waterlvl The folder "Waterlvl" is the preferred location were your water level time-series should be stored. These files can be either in a Microsoft Excel spreadsheet file format (xls) or in a tab-separated values text format (TSV).

Other Files The file "weather_stations.lst" is a resource file that is used to store the results of a weather station search in the Canadian Daily Climate Database (CDCD). The file "graph layout.lst" is also a resource file in which are stored the layout parameters of the well hydrographs that are produced in the hyhdrograph tab of WHAT. The file "weather_datasets_summary.log" is an tab-separated values (TSV) file that contains a summary of all the weather data files contained in the "Input" folder. The file "waterlvl_manual_measurements.xls" is used to input manual water level measurements associated with the water level time-series files stored in the "Waterlvl" folder. These measurements are plotted on the hydrograph and can also be used to adjust the position of the water-level time-series in the vertical axis when the installation depth of the pressure probe is unknown.

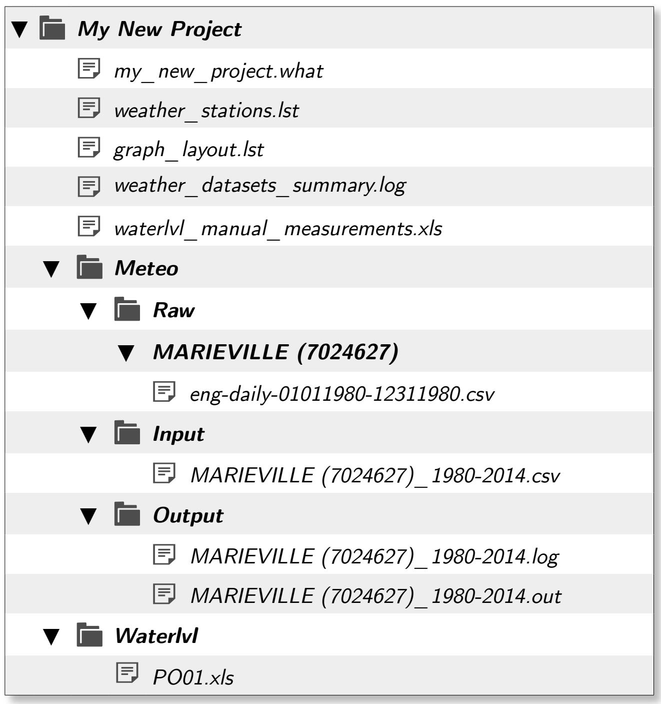


Figure 2.2: Project folder file organization.

3 Creating Gapless Daily Weather Datasets

3.1 Downloading and formatting data from the CDCD

3.1.1 Introduction

The Canadian Daily Climate Database (CDCD) contains daily data for air temperature and precipitation dating back to 1840 to the present for about 8450 stations distributed across Canada. Data can be downloaded manually on the Government of Canada website (`www.climate.weather.gc.ca`) for each year individually and saved in a csv file. This process involves a lot of repetitive manipulations and is a time consuming task. Moreover, the re-organization of the individual data files, saved for each year separately, in a more convenient format can also represent a tedious task when done manually. Alternately, it is possible to order a DVD containing the entire database for a small fee. This option has the disadvantage of only providing an image in time as data cannot be updated.

WHAT alleviates this process by providing a graphical interface to the online CDCD that allows to query for stations interactively using location coordinates, download the available data for the selected weather stations, and automatically organize the data in a format compatible with WHAT. These features are available in the *Download Data* tab shown in Fig. 3.1. This tab consists of a toolbar located at the top of the interface, an area where are displayed the current list of weather stations for which data can be downloaded, and a side-panel to the right where can be manage the formatting of the weather data files that were downloaded for each year individually.

3.1.2 Searching for weather stations

Before any weather data can be downloaded from the online CDCD, a list of stations must first be provided to WHAT. This can be done by selecting an already existing list of stations (files with a "lst" extension) by clicking on the opened document icon located in the toolbar.

Alternatively, it is possible to search for weather stations in the online CDCD by clicking on the magnifying glass icon in the toolbar. This will open a new dialog window (see Fig. 3.2) where it is possible to search for weather stations located within a given radius around a set of location coordinates (latitude and longitude) in decimal degrees. It is possible to further narrow down the search by including only stations with data available within a given period and/or with data available for a minimum number of years.

When all the parameters have been specified, the search for weather stations in the online CDCD can be initiated by clicking on the *Search Stations* button. The resulting stations are automatically displayed in a table, along with information regarding their proximity to the location coordinates used for the search, the years for which data are available, their province, and their climate ID. Selecting stations and clicking on the *Add Stations* button will add them to the current list of weather stations displayed in the *Download Data* tab.

It is possible to remove any weather station from the current list by selecting them and clicking on the toolbar eraser icon. The station list can be saved by clicking on the toolbar floppy disk icon. Alternatively, it is also possible to generate a list of weather stations by creating manually a lst file without using the graphical interface. This is done by retrieving station information from

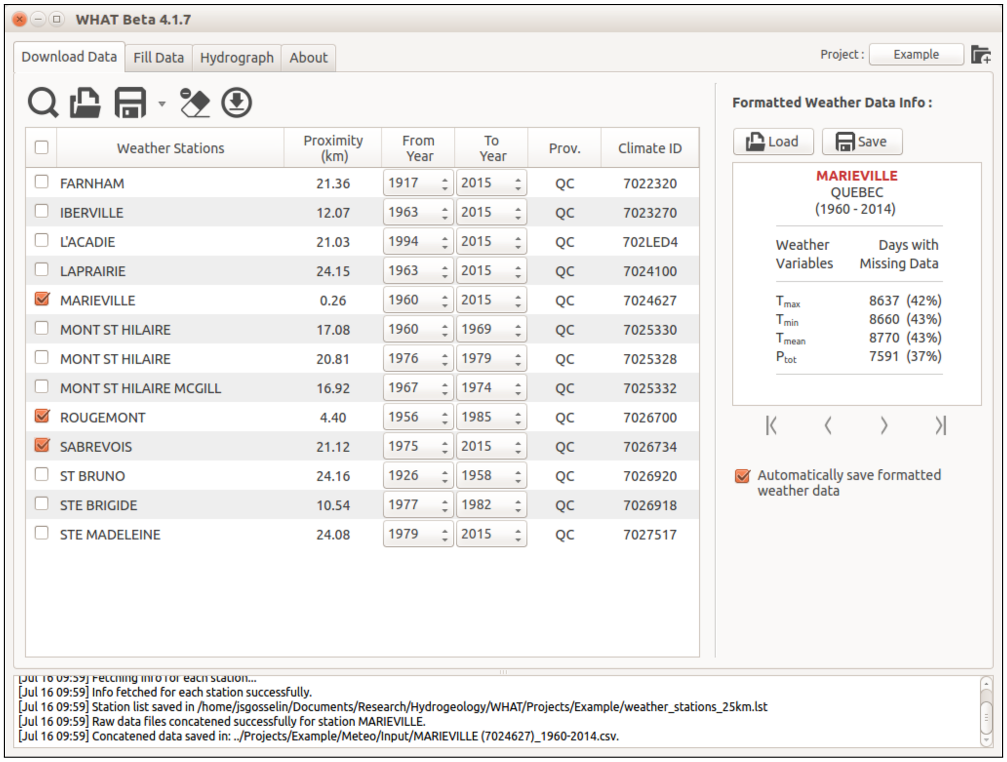


Figure 3.1: Screenshot of WHAT v4.1.7-beta in Ubuntu-Linux 15.04 showing the *Download Data* tab.

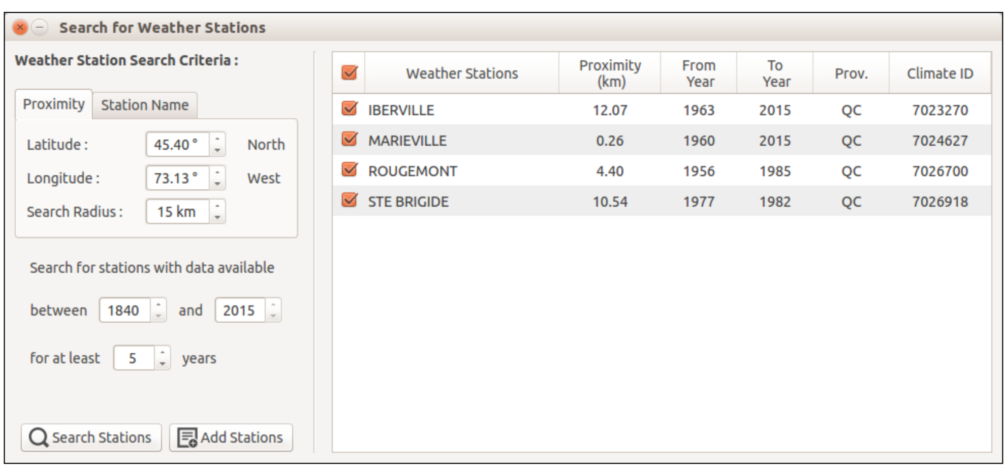


Figure 3.2: Screenshot of WHAT v4.1.7-beta in Ubuntu-Linux 15.04 showing the *graphical interface* to the online CDCD database.

their unique url directly on the government of Canada website (www.climate.weather.gc.ca) and saving the information in a tsv (tabular-separated values) text file with a "lst" extension. A detailed example is presented in Appendix A

3.1.3 Downloading the weather data

Daily weather data can be downloaded from the online CDCD by selecting the desired stations from the list displayed in the *Download Data* tab and clicking on the toolbar icon with the encircled downward arrow. Data will be downloaded for the years specified for each selected station and the results will be saved automatically as a csv (comma-separated values) file in the *Raw* folder (see section 2.4). Weather data for a given station won't be downloaded for the years for which a data file already exist in the *Raw* folder. Detailed information about the downloading process are printed in the console area located at the bottom of the interface (see section 1.4). The downloading process can be stopped at any time by clicking on the stop icon that appears in the toolbar as soon a downloading task is started.

3.1.4 Formatting the weather data

WHAT automatically formats the data as soon as they have been successfully downloaded for a given weather station. To do this, data from each annual file are put together end to end in chronological order. Only the data related to air temperature (mean, max and min) and total precipitation are kept. In addition, days with missing data in the dataset are filled with a NaN (not a number) value. Finally, information on the number of days with missing data for each meteorological variable are displayed in the right side-panel. Alternatively, it is possible to open and format previously downloaded weather data files by clicking on the *Load* button in the right side-panel and selecting the desired files from the dialog window that will open.

By default, WHAT will automatically save the formatted data in a single tsv (tabular-separated values) file in the *Input* folder (see section 2.4). The automatic saving of the formatted data series can be disabled by unchecking the *Automatically save concatenated data* option. From the right side-panel, it is then possible to navigate through the datasets that were formatted over the course of a given session using the left-right arrows and save any dataset manually by clicking on the save button.

3.2 Filling the gaps in daily weather records

3.2.1 Introduction

Climate data are useful in several fields of Earth sciences, including hydrology, hydrogeology and agronomy. However, climate datasets are, most of the time, incomplete. This can represent a major hindrance in various applications, such as for hydrological or hydrogeological model simulations that heavily depend on these data.

Filling the gaps in weather datasets can quickly become a tedious task as the size of the data records and the number of stations increase. Moreover, it can also be rather complex when aspects such as time-efficiency of the method, accuracy of the estimated missing values, and preservation of the statistical properties of the weather time series (probability distribution and normals) are taken into account. This is particularly true for the estimation of missing daily precipitation data because of their high spatial and temporal variability (Simolo et al., 2010).

WHAT addresses this issue by providing an automated, robust, and efficient method to quickly and easily fill the gaps in the daily weather datasets from the files that were produced as described in Section 3.1.4. It is also possible to fill the gaps in weather datasets from files that were not produced with WHAT, provided that the data are formatted in the right format (see Appendix B). In addition, WHAT includes the framework to easily validate and assess the uncertainty of the estimated missing values with a cross-validation resampling technique. These features are available in the *Fill Data* tab shown in Fig. 3.3. This tab consists in a side-panel to the left where the gapfilling procedure can be managed and configured and an area to the right where various outputs are displayed.

3.2.2 Loading the weather data files

When starting WHAT or when a project is opened, the content of the *Input* folder is automatically scanned for weather data files. The results are displayed in a list of weather stations, located under the label *Fill data for weather station*. A summary of the number of days with missing data for each dataset is also produced and displayed in the tab *Missing Data Overview* of the display area, to the right. The icon with the circular arrows, located next to the list of stations, can be clicked to re-scan the *Input* folder for new weather data files to update the list of stations and the summary.

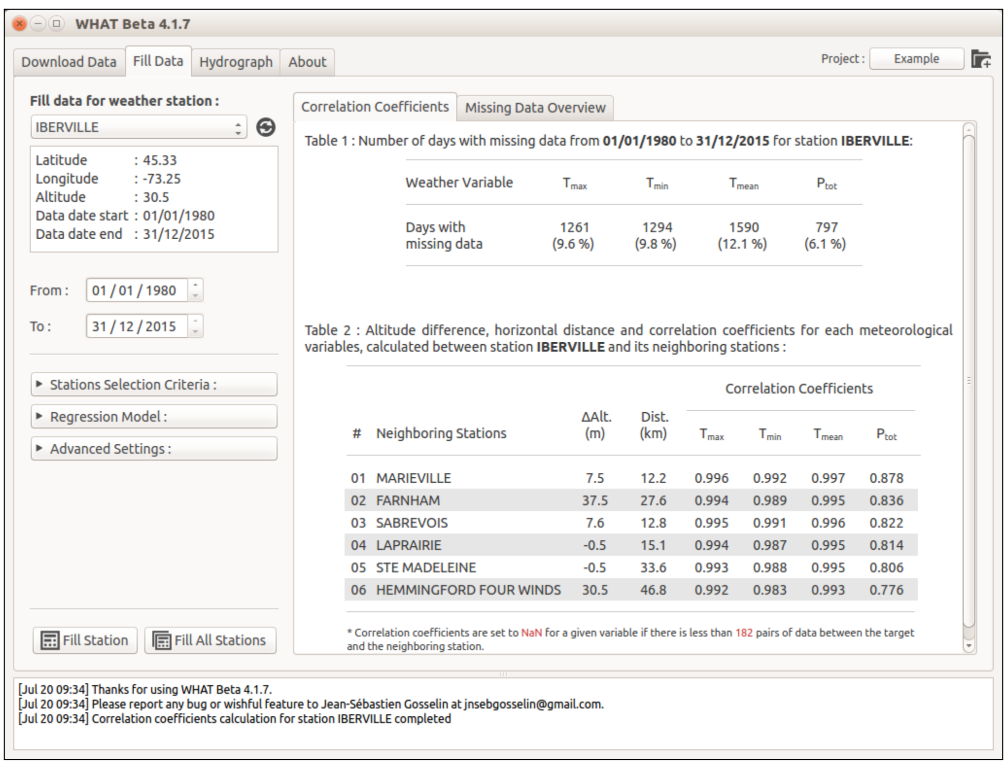


Figure 3.3: Screenshot of WHAT v4.1.7-beta in Ubuntu-Linux 15.04 showing the *Fill Data* tab.

3.2.3 Setting the model parameters

The first step is to select the station for which missing values in the dataset need to be filled from the drop-down list located under the *Fill data for station* section. Under this list are automatically posted information about the currently selected weather station. It is also possible to define the period for which the data of the selected station will be filled by editing the dates in the fields *From* and *To* located in the left side-panel. By default, dates are set as the first and the last date for which data are available for any of the stations of the list.

The method used to estimate the missing data for the selected weather station consists in the generation of a multiple linear regression (MLR) model, using synchronous data from selected neighboring stations from the list. The neighboring stations are selected mainly on the basis of the correlation coefficients computed between their data and those of the selected weather station. The values of these coefficients are automatically displayed in the table located in the right side of the interface when a new weather station is selected from the list. Moreover, among the selected neighboring stations, the ones with the highest correlation coefficients have more weight in the model than those with weak correlation coefficients. For this reason, correlation coefficients that fall below a value of 0.7 are shown in red in the table, as a guidance for the user.

A MLR model is generated for each day for which a data is missing in the dataset of the selected station. This is done because the number of neighboring stations with available data can vary in time. Therefore, for a given date with missing data in the dataset of the selected station, the neighboring stations are selected in decreasing order of their correlation coefficients. Neighboring stations that also have a missing data at this particular date are excluded from the selection process. The maximum number of station that are selected for the generation of the MLR model can be specified in the *Nbr. of stations* field, located in the *Stations Selection Criteria* menu, in the left side-panel. The number of neighboring station that is selected by default is 4. If for a given date, all the neighboring stations have missing data synchronously with the selected station, a NaN value is kept in the dataset at this particular date.

Moreover, the correlation between the data of two stations will, in general, decreases as the distance and the altitude difference between them increase. Therefore, the fields *Max. Distance* and *Max. Elevation Diff.*, both located in the *Stations Selection Criteria* menu, allow to specify thresholds for the distance and altitude difference. Neighboring stations exceeding either one of these thresholds will not be used to fill the gaps in the dataset of the selected station. The default values for the distance and altitude difference are set to 100 km and 350 m, respectively, based on a literature review (Simolo et al., 2010; Tronci et al., 1986; Xia et al., 1999). The horizontal distances and elevation differences calculated between the selected station and its neighbors are shown in the table to the right, alongside the correlation coefficients. The values exceeding its corresponding threshold are shown in red.

Finally, it is also possible to select whether the MLR model is generated using a Ordinary Least Squares (OLS) or a Least Absolute Deviations (LAD) criteria from the *Regression Model* menu, in the left side-panel. A regression based on a LAD is more robust to outliers than a regression based on a OLS, but is more expensive in computation time.

The *Full Error Analysis* option in the *Advanced Settings* menu is discussed in Section 3.2.5

3.2.4 Filling the gaps in the data

The automated procedure to fill the gaps in the dataset of the selected weather station can be started by clicking the button *Fill* located at the bottom of the left side-panel. It is also possible to run this procedure in batch mode to fill the gaps in the datasets of the entire list of weather station by clicking on the button *Fill All Stations*. The parameters for the gap filling procedure will, however, be the same for all the stations.

Once the process is completed for a station, the resulting gapless daily weather dataset is automatically saved in a tsv (tabular-separated values) file with the extension ".out" in the *Output* folder (see Section 2.4). The file is named after the weather station name, climate ID, and first and last year of the dataset. For example, the resulting output file for the station MARIEVILLE in Fig. 3.3 would be "MARIEVILLE (7024627) 1980-2015.out". In addition, detailed information on the values estimated for filling the gaps in the data are saved in a file with the same name as the ".out" file, but with a ".log" extension. Information includes, the names of the neighboring stations, the values of the data used for the estimations, as well as the expected uncertainty of the estimates.

3.2.5 Uncertainty of the estimated values

By default, each time a new MLR model is generated to estimate a missing value in the dataset of the selected station, the model is also used to predict the values in the dataset that are not missing. The accuracy of the MLR model is then approximated by computing a Root-Mean-Square Error (RMSE) between the values estimated with the model and the respective non-missing observations in the dataset of the selected station. The RMSE thus calculated is saved, along with the estimated value, in the ".log" file.

When the *Full Error Analysis* option in the *Advanced Settings* menu is enabled, WHAT will also perform a cross-validation resampling procedure to estimate the accuracy of the model, in addition to fill the gaps in the dataset. More specifically, the procedure consists in estimating alternately a weather data value for each day of the selected station's dataset, even for days for which data are not missing. Before estimating a value for a given day, the corresponding measured data in the dataset of the selected station is temporarily discarded to avoid self-influence of this observation on the generation of the MLR model. The model is then generated and used to estimate a value on this given day and the corresponding observed data is put back in the dataset of the selected station. When a value for every day of the dataset has thus been estimated, the estimated values are saved in a tsv (tabular-separated values) file in the *Output* folder with the extension ".err", along with the ".log" and ".out" files described in Section 3.2.4. The accuracy of the method can then be estimated by computing the RMSE between the estimated weather data and the respective non-missing observations in the original dataset of the selected station. Activating this feature will significantly increase the computation time of the gap filling procedure, especially if the least absolute deviation regression model is selected, but can provide interesting insights on the performance of the procedure for the specific datasets used for a project.

4 Plotting weather and water level data

WHAT make use of the powerful Python package Matplotlib to render the data in publication-quality graph with complex layout. With this tool and the way WHAT is built internally, there is almost no graph configuration that can't be done. The possibility are seldom limited by the UI and the time it requires to implement and design the addition of a new feature in the UI. It is however very fast to make changes in the source code to make your graph exactly the way it is desired. WHAT is buitl in a modular fashion. This means that it is not needed to run the entire program if only a certain feature is needed. For example, it is possible to plus the hydrograph from the Module Hydroprint of WHAT without the UI. If you have any idea, suggestion or request, please contact us. We would like to hear from you. For example, changing the color, adding a legend or plotting multiple water level time series on a same graph is something that can be easily achieved by modifying the source code, but that can take a lot of time to implement in a good and robust UI design.

The tradeoff for the packaging of code into a UI is that the production of graph in the frame of WHAT UI is more strict and allow for less flexibility to the user. Neverthless, it is still possible to produce very good graph from the UI and new options are added frequently to the program to add more flexibility.

The plotting of the weather and water-level data into a same graph in a publication-quality figure is done in the mode “Layout” in the tab called “Hydrograph” (see Figure X).

The mode Layout and Computation both shares the same data. This means that when importing a water level data file or a weather data file in one mode, will affect the content of the other. Thus, the process of importing water level data is the same as the one explained previously in section X. When a water level data file is loaded into memory, the weather data file of the station closest to the well will also be loaded if the folder Output is not empty and will produce a graph with both of these data series. If a weather data file is loaded before a water level data file, only the weather data series will be plotted. It is possible to disable the plotting of either the weather data file or the water-level time series at anytime in order to plot one or the other dataset alone in a single graph. The purpose of this mode is not to explore interactively the data, nor to conduct computation, but to produce publication-quality graph from the data.

The tab hydrograph is equiped of a toolbar at the top and a right panel that is used to edit the graph. By design, it is not possible to interactively modify the content of the graph that is being produced. WHAT display the figure in a bitmap format for performance purposes. The figure however when saved in a pdf or svg format will be fully vectorial for publication-quality work. The figure can be panned by draggin the mouse with the mouse button depressed. Zoom in by pressing the Ctrl key while moving the mouse wheel up and zoom out with mouse wheel down.

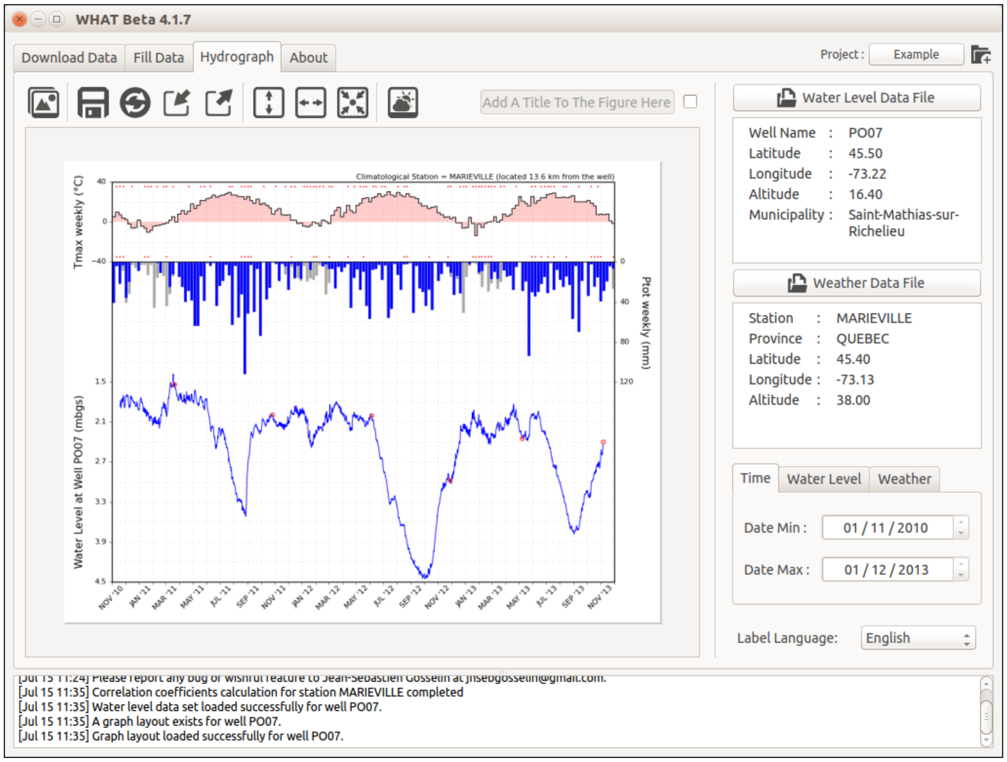


Figure 4.1: Mode “Layout” of the Tab “Hydrograph”.

4.1 Hydrograph Overview

The main feature of this tab is the production of a publication=quality figure that contains both the water-level and weather time-series.

4.2 Toolbar

The toolbar offers various tool that are mostly composed of single action button for automatic formating of the data in the figure. From left to right:

The toggle button is used to switch from one mode to the other. As said previously, both mode shares the same data. Hence, doing manipulation on the dataset in ”Edit” mode will impact what is plotted in the ”Layout” mode. Alternatively, loading a new water level data file will impact both the Edit and Layout mode at the same time. It is possible to switch from one mode to the other at anytime without losing the work that is in progress. Hence, both mode can be used concurrently to edit and analyse the data.

Allow to save the current hydrograph figure either in pdf or svg format. In both format the image is saved in a vectorial format. To convert the figure in a bitmap format such as png, jpg or tiff, software such as acrobat adobe or the very good open source software GIMP can be used. Inkscape is also a very good open source vectorial image editing software that can both work in pdf or svg format.

This button is used mainly for debugging purpose and for experimental features that are not fully integrated in the UI yet. It forces a complete redrawing of the hydrograph.

This button is used to save the current graph layout for future uses. The layout is saved in the file "graph_layout.lst" located in the project folder and is referenced with the ID of the well.

This button is used to force the loading of a previously saved layout associated with the current well. When loading a water-level data file in memory, WHAT will automatically check if there is a layout already saved for that well and will ask the user if he wants or not to load it. If the user refuses, WHAT will try to to a best fit of the data automatically to produce the figure.

The Best Fit buttons are used to force a refitting of the water level data in the vertical axis and in the time axis respectively. This tool does not presently work very well if there is aberrant data in the time-series or if the is a long recuperation curve at the beginning of the dataset following the boring of the well.

This function is used to force WHAT to search for the weather data file in the folder Output of the weather station that is closest to the well, to load it in memory and to plot the results in the hydrograph figure.

4.3 Right Panel

The water level and weather data file section of the Right panel are shared by both mode of the tab hydrograph.

4.4 Water Level Datum

If an appropriate value of the altitude of ground surface at the well location is provided in the header of the input file, it will be possible to switch the datum of the water level when plotting the data from Meters Below Ground Surface (mbgs) to Meters Above Sea Level (masl).

In computation mode however, data are always displayed as meter above ground surface, with the values positive when above ground surface and negative when below. Displaying water-level time-series relative to the ground surface is much more useful than relative to mean sea level or as the height of the water column above the instrument.By displaying the value relative to the ground surface, it is easy to see the width of the unsaturated zone water has to pass through to attained the water table and become groundwater recharge. Depth of the unsaturated zone is a major factor in the delai between the response of the water table to precipitation or snowmelt event and it also play a role in the attenuation of the signal. Moreover, knowing the depth of the water level below the ground surface also gives indication about possible evapotranspiration from the water-table and also flood event.

In Layout mode it is possible to plot the water level relative to four different datum:

- meters above ground surface (mags) - meters below ground surface (mbgs) - meters above logger (mal) - meters above see level (masl) mags and mbgs options have the same reference point, but the vertical axis is inverted in the case of mbgs, with the water level being positive below the ground surface and increasing in value as the depth to the surface increase.

Why Not estimating the precipitation directly at the WELL by interpolation:

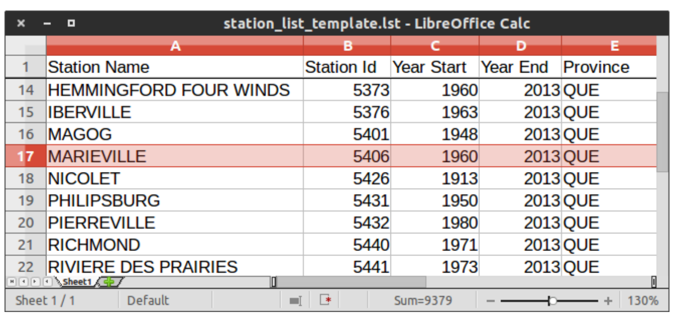
Data are not interpolated to the exact location of the well. It has been decided to keep the original dataset of the station located closest to the well to analyse the data. This is due to the fact that conventional technique for interpoloating weather data tend to surestimate the number of wet day, but underestimate the intensity of stron precipitation event. More advanced and complicated technique are required to circumvent these issues. It has thus been decided that is was prefereable to keep the original data from a single station.

A Creating a Weather Station List Manually

WARNING: this section needs to be updated and revised.

The stations information need to be saved in a tabular-separated values text file with an "lst" extension. A template of a station list (station_list_template.lst) is provided with the program in the Zip archive and an example is presented in Error: Reference source not found. The fields Station Name, Year, Start, Year End and Province do not need to match strictly with the station's URL. These fields can be assigned any name/value by the user and are not directly used in the downloading process of weather data. The only field that is directly used in the download process is Station ID that is a unique number attributed to each weather station.

Once a file containing a list of weather stations' information has been created, it is possible to load it in Rainbird by clicking on the Load button located in the Fetch and Merge tab (Error: Reference source not found). The station list can be refreshed at any time by clicking on the Refresh button.



station_list_template.lst - LibreOffice Calc

	A	B	C	D	E
1	Station Name	Station Id	Year Start	Year End	Province
14	HEMMINGFORD FOUR WINDS	5373	1960	2013	QUE
15	IBERVILLE	5376	1963	2013	QUE
16	MAGOG	5401	1948	2013	QUE
17	MARIEVILLE	5406	1960	2013	QUE
18	NICOLET	5426	1913	2013	QUE
19	PHILIPSBURG	5431	1950	2013	QUE
20	PIERREVILLE	5432	1980	2013	QUE
21	RICHMOND	5440	1971	2013	QUE
22	RIVIERE DES PRAIRIES	5441	1973	2013	QUE

Sheet 1 / 1 Default Sum=9379 130%

Figure A.1: Weather Station List (*.lst) Sample

Step 1 First go to www.climate.weather.gc.ca.

Step 2 At the bottom of the page, click on Advanced Search (see the red arrow in Figure A.2a).

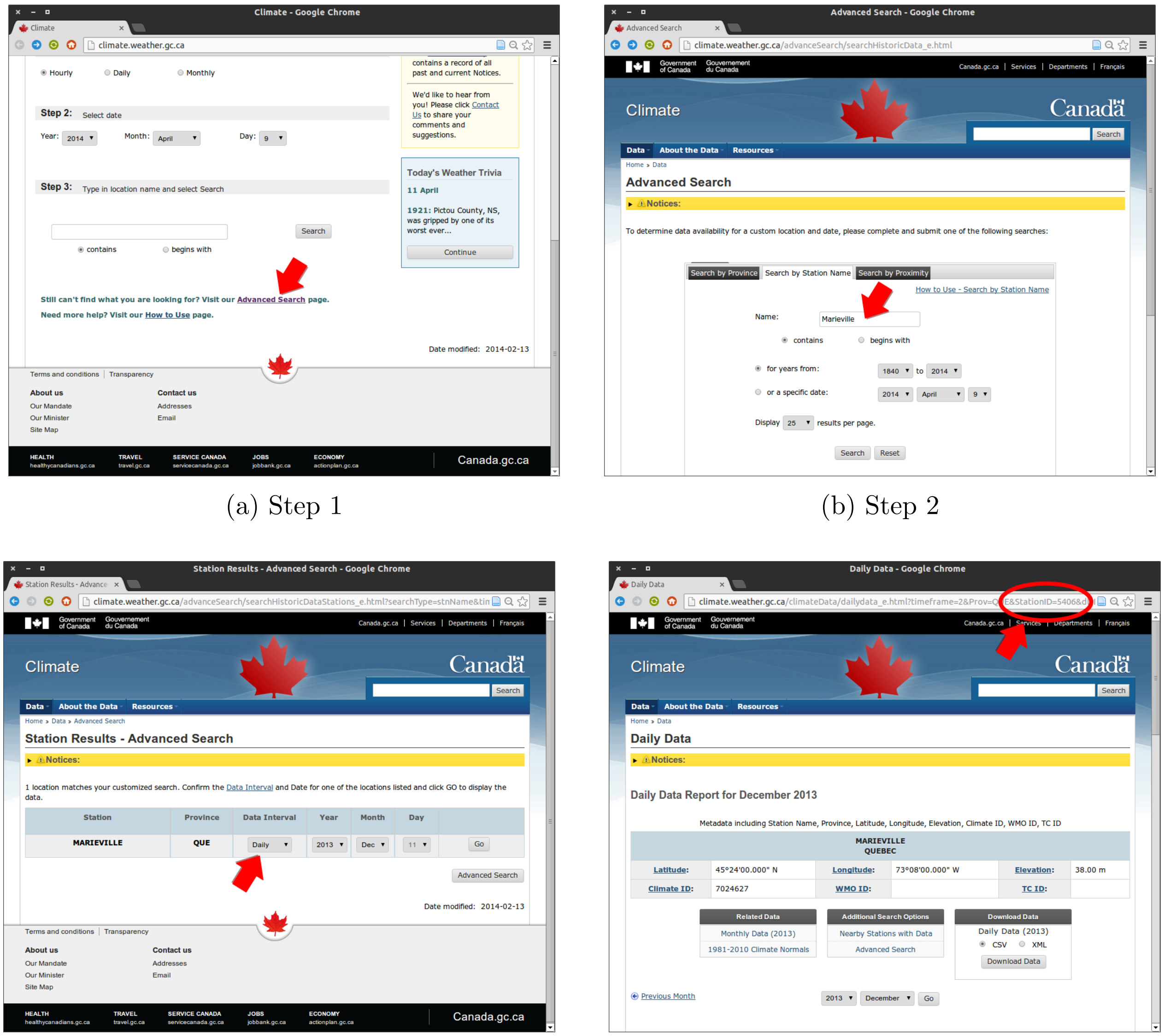
Step 3 Search for a station either by Province, Name or Proximity. For this example, the search was made using the Name of the station (see Figure A2).

Step 4 The research yielded one result (Figure A3). For Data Interval, select Daily and click Go. WARNING: The downloading and formatting of Hourly data is not currently supported in WHAT (but could be in a future version of the software).

Step 5 The URL associated with Marieville weather station is:

http://climate.weather.gc.ca/climateData/dailydata_e.html?timeframe=2&Prov=QUE&StationID=5406&dlyRange=1960-06-01|2013-12-31&Year=2013&Month=12&Day=01.

From this URL, we find that its Station ID is 5406 (see red circle in Figure A4), that is located in the province of Quebec (QUE) and that data are available from 1960 to 2013.



(a) Step 1

(b) Step 2

(c) Step 3

(d) Step 4

Figure A.2: Creation of a custom weather station list how-to

B Description of the weather data file format used in WHAT

It is possible to use weather data from any sources in WHAT, given the right format is used, either to fill the gaps in the weather time series and/or to interpret water level time series. For this purpose, it is recommended to use a copy of one of the sample files that are provided in the project example (distributed with the software) and fill the information and the data directly in it. The file must be kept in a text format using tab-separated values either with the extension ".csv" or ".out", depending if you want to fill the gaps in the weather time series or interpret water level time series. This can be achieved with any standard spreadsheet application such as Microsoft Excel or LibreOffice Calc. The format of the header must be faithfully observed for those files. In addition, "NaN" values must be entered where data are missing. Data must also be in chronological order, but do not need to be continuous over time. That is, missing blocks of data (e.g., several days, months or years) can be completely omitted in the time-series. These missing blocks of data will be filled during the gap filling procedure or will be ignored for the plotting of the hydrograph.

Bibliography

Geoffrey C. Bohling, Wei Jin, and James J. Butler Jr. Kansas Geological Survey Barometric Response Function Software User's Guide. Technical Report Open-file Report No. 2011-10, Kansas Geological Survey, Geohydrology Section, 2011. URL http://www.kgs.ku.edu/HighPlains/OHP/index_program/brf.html. D263.

J J Butler Jr., W Jin, G A Mohammed, and E C Reboulet. New Insights from Well Responses to Fluctuations in Barometric Pressure. *Ground Water*, 49(4):525–533, 2011. ISSN 1745-6584. doi: 10.1111/j.1745-6584.2010.00768.x. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1745-6584.2010.00768.x/abstract. D177.

C Simolo, M Brunetti, M Maugeri, and T Nanni. Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach. *International Journal of Climatology*, 30:1564–1576, 2010. doi: 10.1002/joc.1992. D184.

N. Tronci, F. Molteni, and M. Bozzini. A comparison of local approximation methods for the analysis of meteorological data. *Archives for Meteorology, Geophysics, and Bioclimatology, Series B*, 36(2):189–211, June 1986. ISSN 0066-6424, 1434-4483. doi: 10.1007/BF02278328. URL http://link.springer.com/article/10.1007/BF02278328. D218.

Y Xia, P Fabian, A Stohl, and M Winterhalter. Forest climatology: estimation of missing values for Bavaria, Germany. *Agricultural and Forest Meteorology*, 96:131–144, 1999. doi: S0168-1923(99)00056-8. D182.