1. **Data Management by Projects in WHAT**
   1. **Introduction**

The data are managed in WHAT by project, that is to say, all input and output files relative to a particular project are stored in a common folder called ‘‘project folder’’

This file management system allows to easily backup

projects from one location to the other since all the files related to a given project are saved at the same place.

The first time WHAT is started, the

When you first launch WHAT, the software will automatically open an example with all the necessary files that allows you to easily and quickly test the different functionalities of the program. The title of the current project is shown in the menu bar at the top of the page?. Additional information about the project can be displayed by clicking on the small ‘‘i’’ icon located next to the project name. Only one project at a time can be open in WHAT.

* 1. **Create a New Project**

To start a new project, click on the button *New Project* or click on the small folder icon located at the right end of the WHAT menu bar (see Figure [1](#_bookmark5).1). This will open a new dialog window (see Figure [2.1)](#_bookmark12) where you can enter 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.

Your project information is saved in a file having a ‘‘.what’’ extension. The new folder is created at the location defined by the *Save in Folder* directory path. For example, saving the *My New Project by John Doe* (Figure [2.1](#_bookmark12)) would create a folder named ‘‘My New Project’’ in the directory ‘‘*C://Users/johndoe/WHAT 4.0.5-beta/Projects*’’ and would save 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.

Figure 2.1: New Project dialog window.

* 1. **Open a Project**

To open an existing? project, click on the icon located in the top right corner of the WHAT window. This will open a new dialog window where you can browse your folders to select an existing project file (\*.what), and then click Open. WHAT will then open the project and the displayed project name appearing in the menu bar will change for 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.

* 1. **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 its execution. This file organization is briefly described below and an example is presented in Figure [2.2.](#_bookmark15) The project folder contains two sub-folders named ‘‘Meteo’’ and ‘‘Waterlvl’’ and a few other files.

**Meteo:** The sub-folder *Meteo* contains three sub-?sub-folders named respectively Raw, Input and Output. The binder **Raw** is where the weather data files downloaded from the CDCD are saved. CDCD files contain daily weather data for a given year, each separated with a coma (CSV) for a given year. All the data files for the selected years for a given weather station are saved in WHAT within a common folder named after the station name and its unique identification number (IDN). For example, in Figure [2.2,](#_bookmark15) 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 corresponds to the station IDN.

The binder **Input** contains the formatted weather data files produced from the raw data files, for the selected years of a given station.

These are tab-separated values (TSV) files that are named after the station’s name and IDN.

The binder **Output** is where the gapless weather time-series are saved. Files are produced using the content of the Input folder and a gap-filling procedure. They 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 was estimated by the program to produce the gapless time-series (\*.out files).

**Waterlvl:** The sub-folder ‘‘Waterlvl’’ is the preferred location where 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 hydrograph tab of WHAT. The file ‘‘weather datasets summary.log’’ is a tab-separated values (TSV) file that contains a summary of all the weather data files included in the ‘‘Input’’ folder. The file ‘‘waterlvl manual measurements.xls’’ contains all the available manual water-level measurements from field visits. These values can be added to graphs showing the well hydrograph and hence, they can be used to check and adjust the position of the water-level time-series on the vertical axis when the depth at which the pressure probe was installed in the well is unknown.

Figure 2.2: Project folder file organization.

1. **Water-level time-series preparation**

The process of validation, correcting and updating water-level dataset is not very much complicated, but can represent a fastitious task, especially if there is multiple well installed in the area of study.

WHAT try to alleviate this process by providing tools to easily explore, manipulate and correct the data in a convivial dynamical graphical environment.

WHAT provides a dynamic graphical environment to explore, validate and apply various corrections to water-level time series. This feature is available in the mode ‘‘computation’’ of the tab *Hydrograph* shown in Figure [4.1.](#_bookmark27)

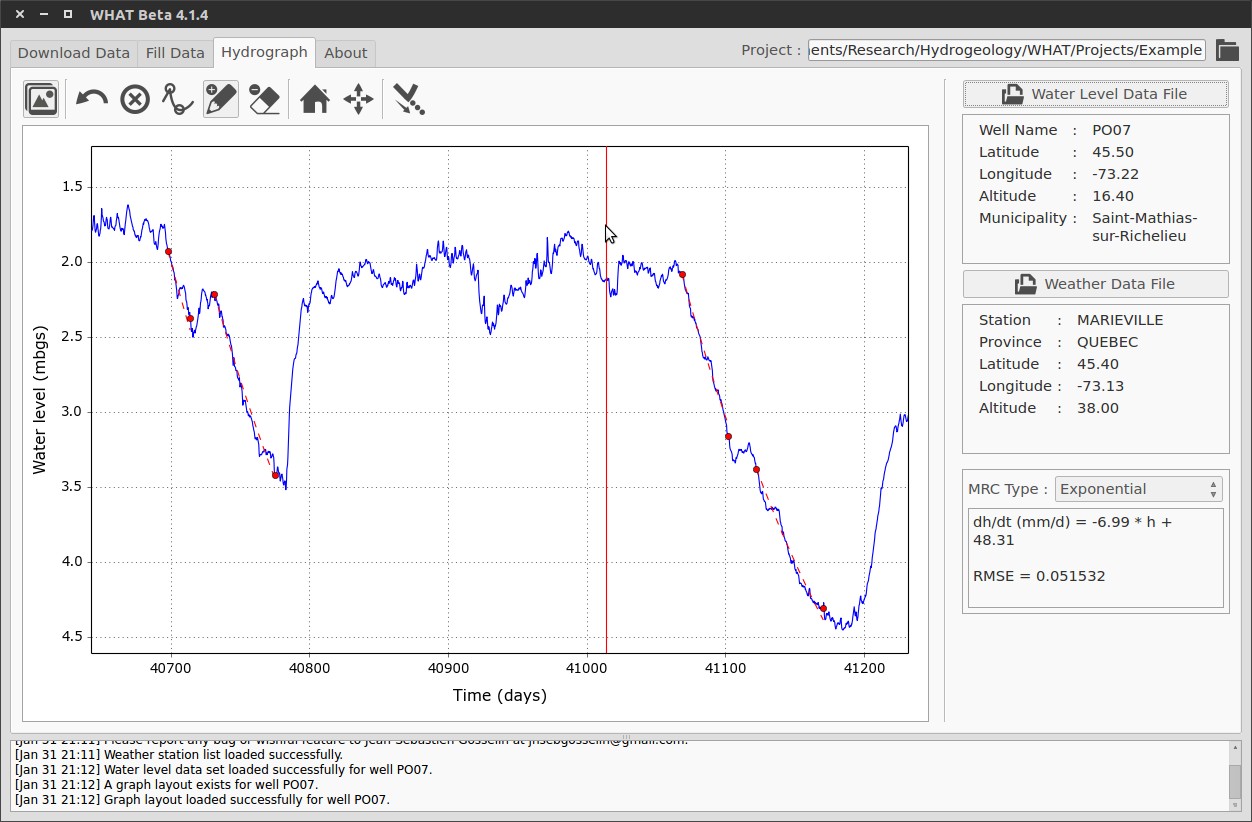


Figure 4.1: Mode ‘‘Computation’’ of the Tab ‘‘Hydrograph’’.

* 1. **Water Level Format**

Water-level data files can be imported in either in a Microsoft Excel 2003 format (.xls) or a tab-separated values (TSV) text file (‘‘.tsv’’). A sample of a water-level data file is provided in the example that is distributed with the software. For the MS Excell format, data must be saved in the first page of the workbook, all the additional pages won’t be read by WHAT for performance purposes.

The water-levels must be entered as the height of the water column above the instrument or the submergence depth. This is generally directly the ouput data of vented water-level data loggers

(gage pressure transducers). However, non-vented devices record the absolute pressure and their output must be compensated for barometric pressure in order to obtain a measure of the water level elevation. This is done by subtracting the barometric record from the water-level record in order to compute the height of the water column above the instrument. This correction must be made before the water-level data are loaded in WHAT. Some software by the data logger manufacturers are able to do this automatically. Alternately, this can be easily done manually when some theory basics are well understood. This is covered in more detail in Section X of Chapter Y.

The measurements must also be accompanied by the times, in the Microsoft Excel numeric time

format, at which they were taken. The following link provides a detailed description and analysis of the numeric time format in Microsoft Excel : <http://www.lexicon.net/sjmachin/xlrd.html>

* 1. **Well Configuration and Location Information**

In addition to the data time-series, various information about the well configuration and location are required in the file header for WHAT to work properly. These information are consists in:

**Well Name** This is an alphanumeric identifier of the well that must be unique to the project. It can contain any combination of alphabetic and numeric characters, but it is recommended to

avoid using the following symbol: %, ’, ”, /, *\*. The Well Name is used to store various information

about the water level time-series (graph layout, data modification, manual measurements, etc.)

and is also used for the generation of the figure labels.

**Latitude and Longitude** Location coordinates of the well in decimal degrees. The well location coordinates are used principally for computing the distance between the well and the weather stations and consequently for the generation of the figure labels.

There exists a great online tool for the conversion of geographic units in various format that is provided by the Montana State University. This tool can be accessed at this web address:

<http://www.rcn.montana.edu/Resources/Converter.aspx>.

**Altitude** Altitude of the ground surface relative to the mean see level at the well location in meters above see level. This value is used to convert the water-level measurements when the datum is changed to mean sea level.

**Municipality** This field is currently not used in WHAT and is for informational purposes only.

**Installation Depth** This is the fixed depth at which water-level data loggers is installed in the well or the piezometer relative to the ground surface. This value is negative if below the ground surface and positive if above.

fixed depth in a well or the piezometer from a stable fixed point called the hanging point, often secured directly to the well casing itself.

* 1. **Manual Measurements**

Water level manual measurements are read automatically from the file named ‘‘waterlvl manual measurements.xls’’ that is located in the project folder (see Section [2.4)](#_bookmark14) when a water level data

file is opened in WHAT.

The information are distributed in 3 columns: the unique ID of the well in which the measurement has been done, the time and the value of the manual measurement. When loading a water-level data file, WHAT will automatically search within this file for every entry that corresponds to the ID of the well.

Manual measurements must be entered in meters relative to the ground surface with the vertical axis positive upward. Measurements below the ground surface are thus negative, and positive when above. Measurement taken relative to the casing of the well must be corrected accordingly.

It is necessary to validate the values taken with the automatic logger with manual measurement on a regular basis.

Long-term monitoring of water levels with the use of automatic data loggers can lead to errors if the data are not validated on a regular basis with manual measurements. [Freeman et al.](#_bookmark91) [(2004)](#_bookmark91) provides a good review on possible causes of errors that may be present in data acquired with submersible pressure transducers.

The convenience and low maintenance of submersible pressure transducers can lead to long intervals between calibration checks and overconfidence in the reliability of the sensor’s data. If checks on the calibration of sensors are not made, data may be erroneous to the point of leading to incorrect hydrologic interpretations.

* 1. **Loading the data and Computation mode overview**

The first step is to open the water-level data file in WHAT. This is done simply by clicking on the *Water Level Data File button* that is located at the top of the right side panel of the tab *Hydrograph*. This will open a new window for the selection of a valide water-level data file. Clicking on select will open the file in WHAT. The water level time-series will then automatically be loaded in WHAT and the data should be automatically plotted. If there is weather data files already present in the ‘‘Output’’ folder, the program will automatically load the file from the closest weather station to the well and will also plot the data along with the water-level measurements.

The correction and adjustment of water-level time-series is done in mode ’Computation’ of the tab hydrograph. By default, this tab is opened in mode Layout. This feature will be covered in detail in Section [5.](#_bookmark36) Switching from the Layout to the Computation mode is done by clicking on the button ‘‘Toggle’’ located at the left end of the toolbar. If a water-level time-series is already imported in WHAT, the data should appear in the graph located in the left pane of the window. If a weather data file has been selected, air temp. and precipitation should also appear on the graph. The display in mode ’Computation’ is dynamic, meaning that it is possible to interact directly with the graph to pan and zoom the content of the graph. By design, only the water-level can be zoomed or panned while the weather data will ajust in the time dimension but will stay static in the vertical axis. This allow a more consistent experience when trying to interpret the water-level

time series.

In order to active the dynamical capability of the graph, the button Pan&Zoom must be toggled on in the toolbar. Panning the vertical or horizontal axes is achieved by holding the left button of the mouse and dragging the mouse horizontally or vertically. Zooming is achieved by holding the mouse right button and dragging the mouse horizontally to zoom in or out in the time axis or vertically for the vertical axis. Zooming both axes equally at the same time can be done by holding the right button and dragging the mouse at an angle of 45 degree toward or away the center of the graph for zooming in and out respectively.

* 1. **Water Level Corrections**

The second step is to apply corrections obtained from field verification measurements. If the record is faulty due to instrumentation or other problems, corrections usually cannot be applied. In general, a missing or faulty record of ground-water level cannot be estimated reliably.

Four types of corrections can be applied to the record:

datum corrections, hung-depth corrections, drift corrections, and calibration corrections. Aber- ration value correction

When data are manipulated in WHAT, the original dataset is not altered directly. Modification are applied to a copy of the originral dataset in order to preserve the later. In addition, each modification applied to a given set of data is registered in a log file.

* + 1. **Aberrant values**

Aberrant values represent water-level measurements that are not representative of the piezometric level of the aquifer in which the well is installed. These values can corresponds to measurement taken when the instrument was out of the water when downloading the data or during a test in the well, or can represent non natural behavior of the level in the well dur for example of pumping during an echantillanage test. Figure Y shows an example of an example with aberrant values that were du to measurement taken while the instrument was out of the water. Tyipically, these measurement will have a value close to zero, since the pressure measured, once corrected for barometric pressure, correspond to a water column of zero height.

These aberrant value complicate the process of interpreting the data and can be removed from the dataset.

Aberrant value can be removed individually in the dataset. To to so, the data can be displayed as dots. The tool remove aberrant data can be used to remove the aberrant data. The process consist in clicking on the button to toggle the edition of data and to hover over the data point that is aberrant. A cross should appear on the data point. Right clicking with the mouse on the data point will remove the points from the data. It is possible to undo up to 10 operation done on the data.

* + 1. **Hung-depth corrections**

Hung-depth errors are caused when the transducer changes relative to its original position, due either to purposeful or accidental raising or lowering of the transducer in the well. In a given project, there may be multiple team that may have access to the well. It is not always possible to track every visit to the well. However, it occurs frequently that the instrument is not replaced exactly at the same place in the well. This generally won’t be perceptible when looking at the hydrograph using a continuous line. However, these errors become apparent when plotting each data point as individual dots. Figure Y show an example of an hydrograph where the water-level data logger was not reinstalled at the same depth in the well dur to mixed cables. The situation was corrected in a subsequent visit to the well.

These occur in the data as discontinuity the variation of the time-series. It is possible to correct these discontinuity in WHAT.

* + 1. **Datum correction**

This operation consist in best-fitting the continuous, aberrant value free, water-level time-series with the manual measurement made in the well. The fit is done by translating and rotatting the curve in order to best fit the manual measurement. The translation allow the correction to any error in the estimation of the depth of installation of the instrument. In addition, if the depth of estallation was previously unknown and set to a value of zero, this will allow the estimation of the depth of installation of the instrument.

the rotating part in the correction is for the correction of any drift that could occur in the automatic measurements of the data by the data logger. The cause and theory for drift in the measurements is covered in detail in ref.

A study of vertical hydraulic head gradients at a well nest in New Hampshire showed that uncorrected data from submersible pressure transducers resulted in an interpretation of reversals in vertical hydraulic-head gradients when none actually occurred (Rosenberry, 1990). In the New Hampshire study, linear adjustment of data based on monthly check measurements would have led to the conclusion that additional water-table fluctuations of up to 0.17 ft occurred when weekly check measurements indicated that sensor drift actually was responsible for those interpreted water-level fluctuations.

So this is a parameter to keep an eye on when doing the validation of the water level time series.

In the project Monteregie Est of the PACES project, the tolerated value for the difference between manual measurement and automatic values was of +- 5 cm.

1. **Plotting the 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).

Figure 5.1: Mode ‘‘Layout’’ of the Tab ‘‘Hydrograph’’.

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.

* 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.

* 1. **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.

* 1. **Right Panel**

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

* 1. **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.