

RECLAMATION

Managing Water in the West

DataDoctor Application (HDB2) User Documentation

Version 3.3

**Excel spreadsheet application created by
Rod Carson and Joe Donnelly, LC Region USBR**

Revision History

Date Updated Application Distributed	Application Updated by	Date Document Updated	Document Version	Document Updated by
8/30/07	Joe Donnelly	9/18/07	3.3	Shana Tighi and Joe Donnelly
2/22/07	Joe Donnelly	3/08/07	3.2	Shana Tighi
9/13/06	Joe Donnelly	9/20/06	3.1	Shana Tighi
6/20/06	Joe Donnelly	6/30/06	3.0	Shana Tighi
10/31/05	Joe Donnelly	11/01/05	2.9	Shana Tighi
7/07/05	Joe Donnelly	7/20/05	2.8	Shana Tighi
6/21/05	Joe Donnelly	6/27/05	2.7	Shana Tighi
5/31/05	Joe Donnelly	6/02/05	2.6	Shana Tighi
3/14/05	Joe Donnelly	3/18/05	2.5	Shana Tighi
	Joe Donnelly	12/27/04	2.4	Shana Tighi
	Joe Donnelly	10/22/04	2.3	Shana Tighi
	Joe Donnelly	9/20/04	2.2	Shana Tighi
	Joe Donnelly	8/26/04	2.1	Shana Tighi
	Joe Donnelly	8/4/04	2.0	Shana Tighi
	Joe Donnelly	6/29/04	1.1	Shana Tighi
	Joe Donnelly	6/25/04	1.0	Shana Tighi

Contents

Revision History	2
Contents	3
i. Summary of Changes since Last Version	3
1. General Information.....	4
1.1 Purpose.....	4
1.2 Versions	4
1.3 Spreadsheet Tabs	4
2. Data Sheet tab	5
2.1 MENU Button.....	5
2.2 FIND Button	6
2.3 GRAF Button.....	7
2.4 REV Button.....	7
3. Main MENU Functions.....	8
3.1 Managing/Updating DataDoctor Parameters.....	8
3.2 Data Queries.....	9
3.2.1 Query Functions.....	9
3.2.2 Query Options	15
3.2.3 Selecting Date/Time Intervals for Analysis.....	16
3.3 After the Query – Creating Graphs and Transferring Data to HDB	20
3.3.1 Graph.....	20
3.3.2 Transfer to HDB	22
3.4 XConnect Edit.....	22
3.5 XConnect Shift Coefficient Scan.....	23
4. RHourInstant Info and RDay Info Tabs.....	25
5. Shift Register Tab	26
6. QuickLook Tab	27
7. Avm Sheet Tab	28
7.1 Avm Sheet Reference Table	28
7.2 Quality Checks for AVM data	29
8. Quality Checks.....	30
9. Updating Changes to HDB2	32
10. Spreadsheet Security	33
11. Glossary of Frequently Used Terms	35

i. Summary of Changes since Last Version

1. Added No Blue Cells option to query userforms
2. Added ability to perform data checks at any time via the main menu
3. Added XConnect Shift Coefficient Scan menu item
4. Added spreadsheet security form – LCHDB username and password are now required to run queries

1. General Information

1.1 Purpose

The DataDoctor application has several primary uses. First, it is used to download and view data from HDB2 in spreadsheet table and graphical formats. Second, it is used to perform quality checks on the data. There are some automatic checks, but it relies to a great extent on manual analysis of and familiarity with the data. When problems are found with the data, the user can manually edit existing data. In addition, the user can manually enter new or missing data, such as phoned-in data values. Lastly, DataDoctor can upload the new or revised data to HDB.

1.2 Versions

There are two versions of the DataDoctor for HDB2 application. They are the DataDoctorShorttermHDB2 application and the DataDoctorLongtermHDB2 application. The capacities of the short term and long term versions are seven years of daily data and four months of hourly data, and one hundred years of daily data and seven years of hourly data, respectively. The short term version is more commonly used, as it runs faster and is thus more practical under typical user conditions. The benefits of using the short term application are speed and computer resource efficiency. The benefits of using the long term application are the greater data capacity. Both versions contain the same spreadsheet tabs and perform the same functions.

1.3 Spreadsheet Tabs

The table below lists the name of each spreadsheet tab in the DataDoctor application and a general description of the purpose it serves for the application:

Tab Name	General Description/Purpose
data sheet	Primary viewing/editing sheet. Macros update data as specified and perform basic quality checks
Graph	Displays a single line graph as specified by the user
Multiple Graph	Displays a line graph with up to five data series on the same set of axes as specified by the user
Two Axes Graph	Displays a line graph with two data series and two y-axes
Multiple Graph Data	Contains the data used to create the multiple graph. Only populated when Multiple Graph is selected from the menu.
Shift Register	Contains a lookup table, organized by Site, RHourInstantID, and RDayID. The lookup table contains dates and shift values for use in quality checks on HDB shift data values.
RHourInstant Info	Reference sheet and quality check parameters for hour and instant data listed by site abbreviation and data type of interest.*
RDay Info	Reference sheet and quality check parameters for daily data

	listed by site abbreviation and data type of interest.*
QuickLook	Lists site_datatypes to be included in Quick Look data display options
XConnect Edit	Used for quality control, indicates which data parameters have missing or invalid data and the dates/times where data is bad
Lake Contents	Contains selections from the elevation-volume tables for Lakes Mead, Mohave, and Havasu
Avm Sheet	Reference sheet and quality check parameters for data pertaining to acoustic velocity meters (aka acusonic flow meters) listed by site abbreviation and data type of interest. *
Array Sheet	Contains values needed by macros to perform calculations and other functions. Values are updated automatically by macros as specified.
Miscellaneous Data	This tab will be used in future versions of the DataDoctor in editing telemetry data

* In HDB2 the site and data type combination is called the site_datatype.

2. Data Sheet tab

The Data Sheet tab is the main worksheet in the DataDoctor application. All major application functions are initiated from this sheet, and all manual data editing is performed here. There are four buttons on the top left corner. The following sections describe how each button may be used.

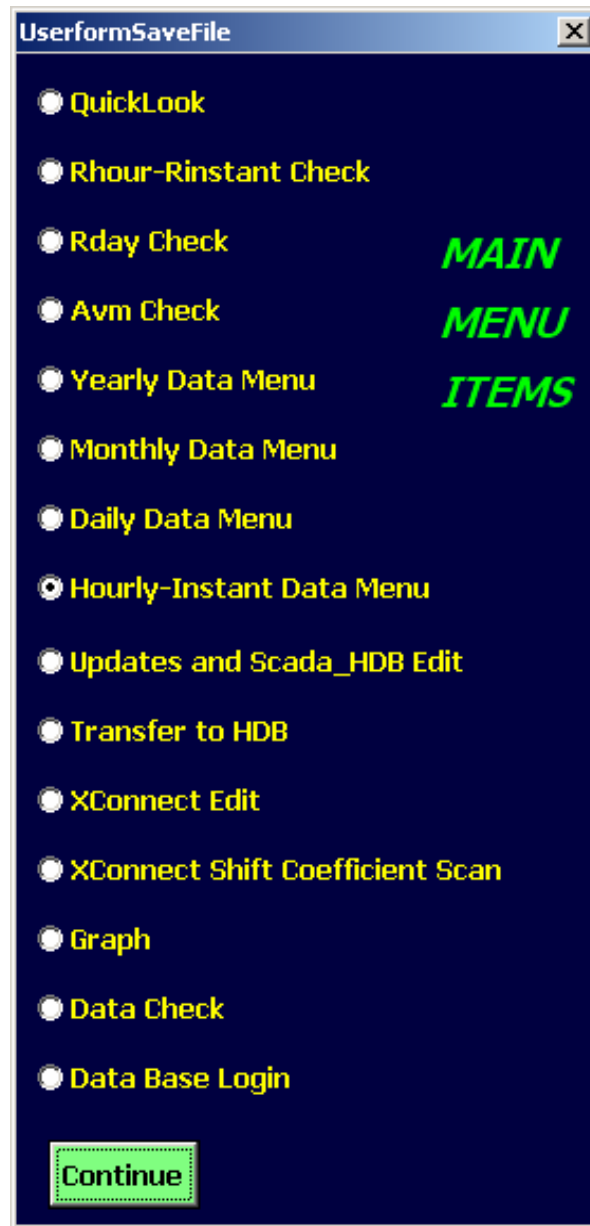
	A	B	C
1	MENU	HDMLC	HDMLC
2	FIND	fbag	qd
3	GRAF		
4	Source = Interval		
5	Repeat (long)		
6	Repeat (sets)		
7	Missing Data Errors		
8	Outside Limits Errors		
9	Rate of Change Errors		
10	Max ROC Error		
11	ROC Isolation		
12	End Time Values		
13	6/30/2006	1128.25	12475
14	6/30/2006 1:00	1128.26	13141
15	6/30/2006 2:00	1128.26	5893
16	6/30/2006 3:00	1128.26	5663

Buttons on top left corner – MENU, FIND, GRAF, REV

2.1 MENU Button

This button is the primary mechanism for initiating any of the main DataDoctor actions, such as querying data or creating a time series graph. When selected, a menu box pops up

titled Main Menu Items. Select the action you want DataDoctor to perform and select the Continue button.



Main Menu Screenshot

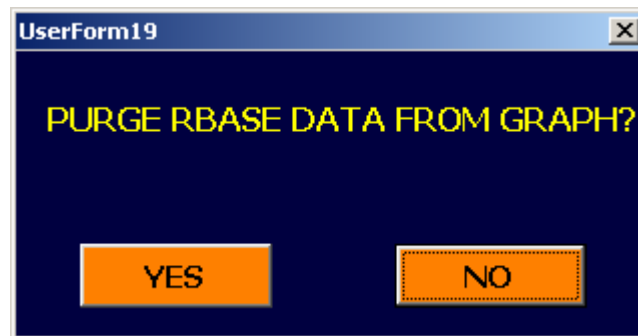
2.2 FIND Button

This button is used after a query has been completed and basic quality checks have been performed. Selecting the FIND button moves the cursor to the next cell with an error of any type. The cursor starts at the cell that is currently selected and moves along the

columns from top to bottom, left to right. Once the cursor has reached the bottom of the last (far right) row of data, it wraps around to the top of the first (far left) row.

2.3 GRAF Button

This shortcut button will create a time series line graph for a single data series. Select a cell anywhere in the column of data you wish to graph and select the GRAF button. A box will pop up asking if you want to exclude data that appears in blue cells from the plot. More explanation about purging bogus and missing data from graphs can be found in Section 3.3.1.



Screenshot of Box Asking if You Want to Exclude Un-derived Data from Graph

2.4 REV Button

The REV button is used while editing data values. Whenever a user enters a value into a cell, either as a new entry or a change to a currently existing value, the background color of the cell changes to red. The red background indicates to DataDoctor that the data value has been updated and should be transferred to HDB2 the next time the transfer function is called. If a user enters a new data value, but then later decides that the value should not be updated to the database, he/she can select the cell in question and click on the REV Button. The background color of the cell will change to white, thereby removing the trigger that tells DataDoctor to transfer the value to HDB.

Note: When you select the REV button, the cell content will NOT revert back to the original, pre-edited value. Only the background color will change.

3. Main MENU Functions

3.1 *Managing/Updating DataDoctor Parameters*

Updates and Scada_HDB Edit

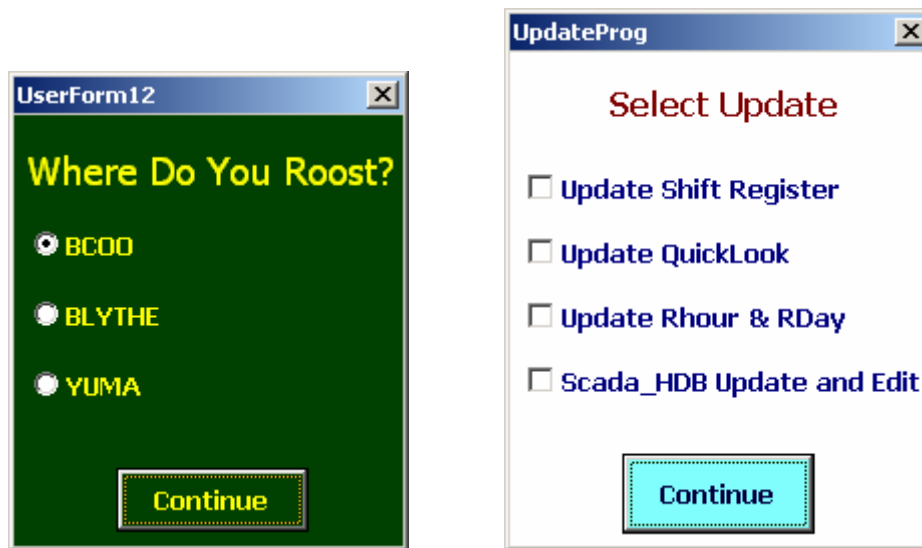
The Updates and Scada_HDB Edit menu item does several tasks, and is customized to function correctly at several different office locations within the Lower Colorado Region. The tasks that can be accomplished are:

- 1) Update the Shift Register worksheet from the master file or a copy of the master file on a local directory,
- 2) Update the QuickLook worksheet from a user's local directory or the master directory,
- 3) Update the RHourInstant Info and RDay Info worksheets from a user's local directory or the master directory, and
- 4) Copy and paste edited data from the SCADA_HDB file to the DataDoctor data sheet

Users should keep a copy of the Shift Register file and a copy of their personal QuickLook file on their personal drive. At BCOO, users should store the files in their D:\DataDoctor\ directory. The master Shift Register file is located at BCOO directory G:\COM460\Water Scheduling Documents\Shift Register.xls. This is the file that gets manually updated when new values are called in. (See Section 5 on Shift Register Tab below)

When Updates and Scada_HDB Edit is selected from the menu, the first dialogue box that pops up asks you to select your office site. The offices currently supported by DataDoctor are BCOO, Blythe, and Yuma. It is necessary to specify your location because the local file directory structure is different at each Reclamation site. In order to retrieve data from files stored on a local directory, DataDoctor needs to know which file structure it is dealing with.

Once you have specified your office location, another dialogue box pops up asking you which task you want to perform. Users at any site can update the Shift Register or QuickLook worksheets. However, only users at the BCOO site can edit data from the SCADA_HDB files. An error message will appear if users from any other site select this task.



Screenshots of Updates and Scada_HDB Edit Menu Boxes

If you want to update the shift register, or the RHourInstant Info and RDay Info worksheets, a third dialogue box will pop up asking you to specify which source file you want to use – the master copy from the G: drive or your copy on your local (D:\) drive. DataDoctor then selects the appropriate file, copies the contents of the file, and pastes them onto the Shift Register tab or RHourInstant Info and RDay Info tabs as applicable.

If you want to update the QuickLook worksheet, no additional dialogue boxes will appear. DataDoctor will select the local QuickLook file, copy the contents of the file, and paste them onto the QuickLook tab.

The SCADA_HDB Update and Edit option is used in conjunction with the SCADA_HDB editing application called SCADA_HDB EDITING SHEET.xls. The SCADA_HDB EDITING SHEET.xls application is used to load and format ASCII text files sent from Hoover IT personnel. Then the SCADA_HDB Update and Edit option in the DataDoctor application is used to load the data into HDB. Additional documentation on this topic can be found in the SCADA_HDB EDITING SHEET.xls file.

3.2 Data Queries

3.2.1 Query Functions

QuickLook

QuickLook should be selected when you want to view and edit selected site_datatypes from a predefined list. When QuickLook is selected, a second dialogue box pops up titled “QuickLook Check”. The box contains a list of QuickLook options as well as Beginning Date and Ending Date spaces. The default value for Beginning Date is the current day,

midnight (the time is not displayed but should be understood as midnight). The default value for Ending Date is the current day, current time. Users may click on Beginning Date to select more predefined options. Whenever you click the beginning date text box today minus one day is automatically selected. If you want to see data over a time period not listed, you can manually enter a beginning data and an ending date. See Section 3.2.3 below for more details on selecting a period of record.

The user may also specify if DataDoctor should retrieve data from the r_base table or the interval tables (r_instant and r_hour). By default, DataDoctor will retrieve data from the interval tables. See Section 3.2.2 below for more details on selecting the source data.

Lastly, the user must specify a QuickLook option to display. The following QuickLook options have been predefined and distributed with updated DataDoctor software:

- Standard
- Dams
- Hoover Avms
- Shifts
- Red Mtn
- Xmas
- Metal Mtn
- Black Pt.
- Tgp
- New Sites
- Blythe Gages
- SCADA_HDB
- Yuma
- XConnect
- Water Qlty
- Wtrtemps
- USGS Sites
- Elev Sites

Individual users may customize and add to the QuickLook options to suit their needs. See section 5 below for more details on customizing and updating Quick Look menu options.

After the date/time interval is entered and a QuickLook option is selected, DataDoctor populates the data sheet tab with all sites and parameters specified in the QuickLook tab, under the column heading selected in the dialogue box. Data is queried from HDB2 and populates the columns. By default, missing values are identified with blue highlighted cells. However, users may specify that missing values should not have special formatting by selecting the “No Blue Cells” box. Section 3.2.2 below has more information on the “No Blue Cells” option.

Once the data sheet has been updated, another dialogue box will pop up asking if you want to perform quality checks on the data. If selected, DataDoctor will perform quality checks using data specified in the RHourInstant Info tab. See quality checks Section 6 for more details.

Screenshot of QuickLook Menu Box

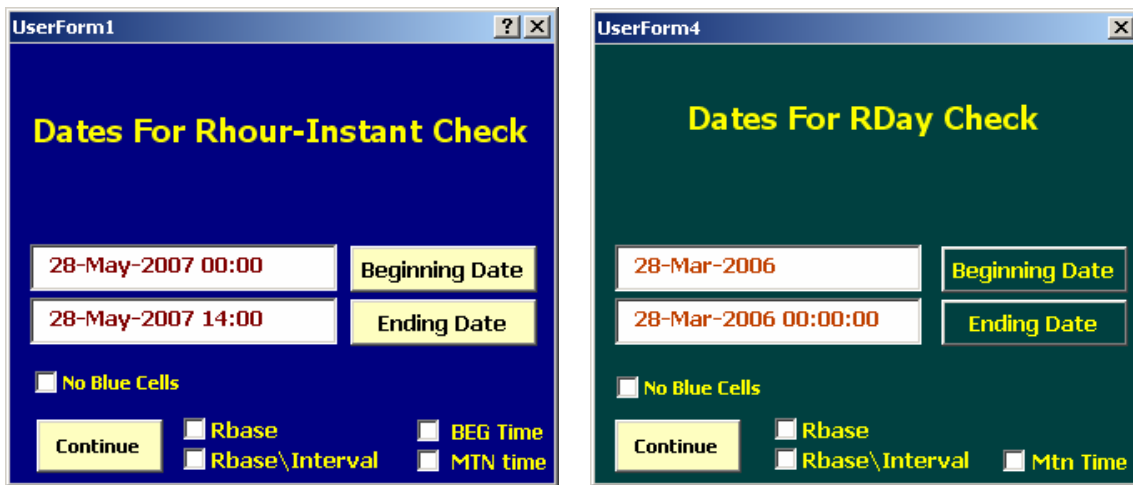
Rhour-Rinstant Check

Rhour-Rinstant Check should be selected when you want to view and edit hourly data from all sites listed in the RHourInstant Info tab not marked with an Omit indicator. When Rhour-Rinstant Check is selected, a second dialogue box pops up titled “Dates For Rhour-Rinstant Check”. The user can enter a Beginning Date and an Ending Date in the spaces provided or click on either the Beginning Date or Ending Date buttons to select a date/time interval. See Section 3.2.3 below for more details on the date selection dialogue box.

The user may also specify if DataDoctor should retrieve data from the r_base table or the interval tables (r_instant and r_hour). By default, DataDoctor will retrieve data from the interval tables. See Section 3.2.2 below for more details on selecting the source data.

After the date/time interval is entered and the continue box selected, DataDoctor populates the data sheet tab with all sites and parameters in the RHourInstant Info tab, in the order specified, as entered under the ‘Analysis Order’ column (excluding site_datatypes marked with the Omit indicator). Data is queried from HDB and populates the columns. As with the QuickLook menu item, missing values are identified with blue highlighted cells unless the user selects the “No Blue Cells” box.

Once the data sheet has been updated, another dialogue box will pop up asking if you want to perform quality checks on the data. If selected, DataDoctor will perform quality checks using data specified in the RHourInstant Info tab. See Section 8 for more details on quality checks.



Screenshots of RHourInstant Menu Box and RDay Menu Box

RDay Check

RDay Check should be selected when you want to view and edit daily data from all sites listed in the RDay Info tab not marked with an Omit indicator. When RDay Check is selected, a second dialogue box pops up titled “Dates For RDay Check”. See Section 3.2.3 below for more details on the date selection dialogue box.

The user may also specify if DataDoctor should retrieve data from the r_base table or the interval table (r_day). By default, DataDoctor will retrieve data from the interval table. See Section 3.2.2 below for more details on selecting the source data.

After the date/time interval is entered and the continue box selected, DataDoctor populates the data sheet tab with all sites and parameters in the RDay Info tab, in the order specified, as entered under the ‘Analysis Order’ column (excluding site_datatypes marked with the Omit indicator). Data is queried from HDB2 and populates the columns. Missing values are identified with blue highlighted cells unless the user selects the “No Blue Cells” box.

Once the data sheet has been updated, another dialogue box will pop up asking if you want to perform quality checks on the data. If selected, DataDoctor will perform quality checks using data specified in the RDay Info tab. See Section 6 for more details on quality checks.

Avm Check

Avm Check should be selected when you want to view and edit data from AVM sites (currently at Davis and Parker only). When Avm Check is selected, a second dialogue box pops up titled “Dates For Avm Check”. See Section 3.2.3 below for more details on the date selection dialogue box.

After the date/time interval is entered and the continue box selected, DataDoctor populates the data sheet tab with all sites and parameters in the Avm Sheet tab, in the order specified, as entered under the 'Analysis Order' column. Data is queried from HDB2 and populates the columns.

Once the spreadsheet has been updated with fresh AVM data from HDB, three new buttons appear on the data sheet tab and the appearance of the top rows change. In the top left corner, next to the MENU button is a button titled RED. Along row 10, are two buttons titled PASTE. These buttons can be used for quality checks and editing. See Avm Sheet Tab Section 7 for details.

Yearly Data Menu

When Yearly Data Menu is selected, a second dialogue box pops up titled "Select Sites". In addition to spaces and buttons to use for entering a Beginning Date and an Ending Date, there is also a list of all site_datatypes. You can select multiple site_datatypes by clicking on them with the mouse. You do not need to hold down the Shift key or the Ctrl key. Up to 20 site_datatypes may be selected for a single data check.

Once the Continue box selected, DataDoctor populates the data sheet tab with all site_datatypes selected in the dialogue box. Selections in the dialogue box are placed in alphabetical order. Missing values are identified with blue highlighted cells unless the user selects the "No Blue Cells" box.

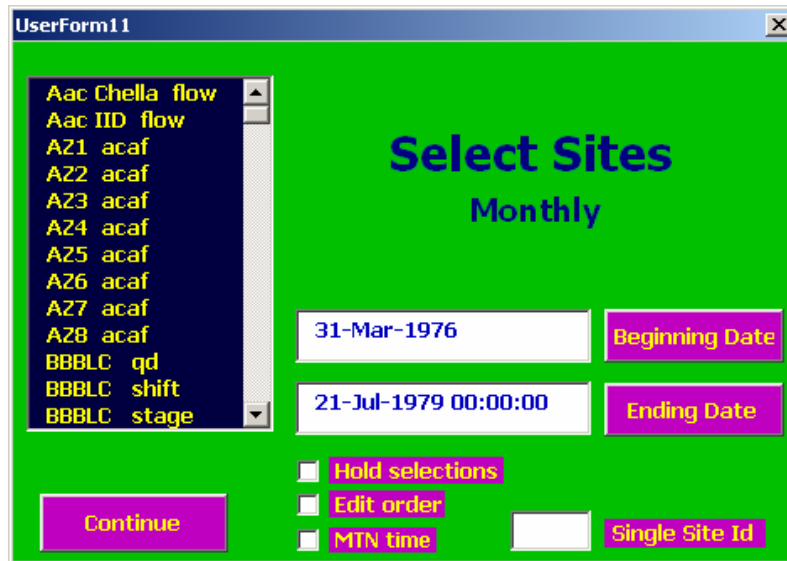
Note: Users cannot edit or perform automated quality checks for yearly data. DataDoctor can only be used to view yearly data.

Monthly Data Menu

When Monthly Data Menu is selected, a second dialogue box pops up titled "Select Sites". In addition to spaces and buttons to use for entering a Beginning Date and an Ending Date, there is also a list of all site_datatypes. You can select multiple site_datatypes by clicking on them with the mouse. You do not need to hold down the Shift key or the Ctrl key. Up to 20 site_datatypes may be selected for a single data check.

Once the Continue box selected, DataDoctor populates the data sheet tab with all site_datatypes selected in the dialogue box. Selections in the dialogue box are placed in alphabetical order. Missing values are identified with blue highlighted cells unless the user selects the "No Blue Cells" box.

Note: Users cannot edit or perform quality checks for monthly data. DataDoctor can only be used to query and view monthly data.



Screenshot of Monthly Data Menu Box

Daily Data Menu

When Daily Data Menu is selected, a second dialogue box pops up titled “Select Sites”. In addition to spaces and buttons to use for entering a Beginning Date and an Ending Date, there is also a list of all site_datatypes. You can select multiple site_datatypes by clicking on them with the mouse. You do not need to hold down the Shift key or the Ctrl key. Up to 20 site_datatypes may be selected for a single data check.

Once the Continue box selected, DataDoctor populates the data sheet tab with all site_datatypes selected in the dialogue box. Selections in the dialogue box are placed in alphabetical order. Missing values are identified with blue highlighted cells unless the user selects the “No Blue Cells” box.

Once the data sheet has been updated, another dialogue box will pop up asking if you want to perform quality checks on the data. If selected, DataDoctor will perform quality checks as specified in the RDay Info tab.

Hourly-Instant Data Menu

When Hourly-Instant Data Menu is selected, a second dialogue box pops up titled “Select Sites”. In addition to spaces and buttons for use in entering a Beginning Date/Time and an Ending Date/Time, there is also a list of all site_datatypes. You can select multiple site_datatypes by clicking on them with the mouse. You do not need to hold down the Shift key or the Ctrl key. Up to 20 site_datatypes may be selected for a single data check.

Once the Continue box is selected, DataDoctor populates the data sheet tab with all site_datatypes selected in the dialogue box. Selections in the dialogue box are placed in alphabetical order. Missing values are identified with blue highlighted cells unless the user selects the “No Blue Cells” box.

Once the data sheet has been updated, another dialogue box will pop up asking if you want to perform quality checks on the data. If selected, DataDoctor will perform quality checks as specified in the RHourInstant Info tab.

Screenshot of RHourInstant Menu Box

3.2.2 Query Options

Source Data

By default, DataDoctor will retrieve data from the interval tables; that is, from *r_instant*, *r_hour*, *r_day*, *r_month*, or *r_year*. However, when querying instant, hourly, and some daily data, the user may specify that DataDoctor should retrieve data from the *r_base* table. When the data selection box pops up, there are two small toggle options along the bottom of the dialogue box.

If the user checks the *Rbase* box, the data will be queried directly from the *r_base* table. If the user checks the *Rbase\Interval* box, the data will first be queried from the appropriate interval table. As always, any cells missing data will appear light blue. When the retrieval from the interval table is complete, DataDoctor will query the *r_base* table for data that was missing from the interval table. Any data in *r_base* that did not get propagated to the interval tables will appear in dark blue text. The background will remain light blue to indicate that the data was missing in the interval table.

The source of the data (*r_base*, interval, or both) will be specified in cell A3 of the data sheet.

No Blue Cells

By default, any cells on the data sheet tab that have missing or invalid values are formatted with a light blue background and dark border. If the “No Blue Cells” checkbox is selected and DataDoctor finds data time slots containing missing or invalid data, the cell background of the affected time slot cell will retain a white background color and any values placed therein will have a blue font.

The advantage of using the default option (“No Blue Cells” box left unchecked) is that cells with a light blue background tend to be easier to see and give data editors a greater awareness of the existence and magnitude of missing or bogus data cells.

The advantages to selecting No Blue Cells are query processing speed and memory retention. No Blue Cell queries run a bit faster because the program does not have to take time to place a light blue background and a dark border on missing or bogus cells. Less memory is also required if light blue backgrounds and dark borders are not placed.

The advantages inherent in the No Blue Cells selection are maximized when queries involving long time periods or large numbers of missing or invalid data cells. The advantages are barely noticeable on the other hand when queries involving short time periods or minimal bogus or missing data are made.

Hold Selections and Edit Order

The Hold Selections box and the Edit Order box appear as options for the Yearly, Monthly, Daily, and Hourly Data menu items. The default option for both boxes is unchecked, or inactivated.

By default, if a query has completed and the user wants to run another query with the same parameters, the user must select all data parameters again when he or she brings up the query menu dialogue. When the user selects the Hold Selections box, DataDoctor will remember the parameters selected in the current query. The next time the user brings up the menu dialogue, the same parameters will be selected.

The default order for queried data to appear on the data sheet is alphabetical order, from left to right. When the Edit Order box is checked, queried data is printed on the data sheet in the order specified on the RHourInstant Info tab or RDay Info Tab. The edit order must be specified by the user by entering numbers into Column D, labeled Edit Order, of the Info tabs prior to activating the Edit Order option on the menu dialogues. For convenience, a default edit order has been entered by the programmer; however each user may customize the data appearance order to meet his or her individual needs.

3.2.3 Selecting Date/Time Intervals for Analysis

Whenever you select a data check option from the menu, a dialogue box will pop up prompting you for a beginning date and an ending date. Only the QuickLook option has

default dates. The default begin date is the current day, midnight, and the default ending date is current day, current time.

For QuickLook checks, you can change the default dates to include more historical data by clicking on either the Beginning Date or Ending Date buttons. Date menu options range from today minus one day through today minus six days, as well as the most recent one to three weeks, the most recent one to three months, and calendar year quarters. The largest period of time a QuickLook will cover is three months unless you manually type in the dates into the spaces provided. The QuickLook time range options are based on a beginning time equal to the current date. The only exceptions are the quarterly time ranges, for which you can query the current year or the previous year. There are several things to keep in mind when running a quarterly QuickLook query:

- If the current date is prior to the time range of the quarter you select, you will get an error message telling you that you selected a quarter that does not exist yet and to hit OK to return to the QuickLook menu.
- If the current date is within the time range of the quarter you select, you will get a message telling you that the query will continue but your data listing will include blue squares for dates that have not occurred yet. You also have the option of quitting the query and returning to the QuickLook menu.
- If the current date is after the last day included in the quarter, the program will query the current year for the quarter selected.
- If you want to select a quarter from the previous year, you must check the “QTR X LAST YEAR” Box in addition to selecting the calendar quarter.



QuickLook Date Selection Dialogue

The screenshot shows a window titled "UserForm10" with a dark red background. It contains two sections for date selection. The first section, "Beginning Date", has three dropdown menus for day (28), month (Mar), and year (2006), followed by a "Hour" dropdown set to 0000. The second section, "Ending Date", has identical dropdown menus. Below these are six buttons: "Go" (yellow), "Today" (light blue), "Today & Yesterday" (light blue), "Last 7 Days" (light blue with a dashed border), "Last Month" (light blue), and "Year to Date" (light blue).

Date Selection Box for All Other Queries

Another unique element of the QuickLook selection box is the Time Offset checkbox. The purpose of the Time Offset is to allow users to perform QuickLook queries before the automated top-of-the-hour data transfers have completed. The default value in the checkbox is 14 minutes. This is primarily for cosmetic reasons – some data gets loaded from telemetry 3 minutes after the hour, some at 7 minutes after the hour, etc. By 15 minutes after the hour, all hourly data is loaded into the database; therefore, leaving the time offset at 14 insures that the last row of data does not have any artificial blue cells, which may lead to confusion by the data editing personnel. There are situations in which it is important to see what data has been transferred to the interval or r_base tables prior to 15 minutes past the hour. In such situations the user can enter a lower offset, as low as 0, to view data as it gets loaded into HDB just a few minutes past the hour.

The following date/time functionality is used for all other (non-QuickLook) data checks:

There are two ways to change the beginning and end dates. First, you can type in the desired dates manually in the spaces provided. If you type in the dates, DataDoctor will not store the dates you entered and you will have to re-enter the dates the next time you run a data check during the DataDoctor user session. The second method of changing the date/time interval is to click on the Beginning Date or Ending Date buttons. Clicking either of these buttons will bring you to the date selection menu box shown above.

You can select a beginning date and ending date by selecting options from drop down menus for day of month, month, year, and hour. Once you select your options from the drop down menus, click the Go button to return to the previous dialogue box.

For daily data and hourly/instant data, the following predefined options are also included on the date selection menu:

- Today – will automatically populate the beginning date with the current date, midnight and the ending date with the current date, current hour
- Today & Yesterday – will populate the beginning date with the previous date, midnight and the ending date with the current date, current hour
- Last 7 Days – will populate the beginning date with the date from seven days ago, midnight and the ending date with the current date, current hour
- Last Month – will populate the beginning date with the date of the same day of month, previous month, midnight and ending date with the current date, current hour.
- Year to Date – will populate the beginning date with January 1 of the current year and the ending date with the current date.

Selecting any of these four options will return you to the original dialogue box with the dates changed as listed above. If you select a date/time interval from the date selection menu, (using either the drop down menus or the predefined selection buttons) the date values will be stored in DataDoctor for the remainder of the user session. You must save the DataDoctor application document in order to save the dates for the next user session.

BEG time

All data in HDB have a start date/time and an end date/time to specify the interval over which the record spans. DataDoctor has only one column for date/time; therefore the user must specify if the data should be listed with the starting or ending date/time. By default, DataDoctor lists the end date/time values whenever a query is run. To list the start date/time values, check the BEGtime box on the date selection dialogue.

The type of date/time value (start or end) will be specified in cell A11 of the data sheet.

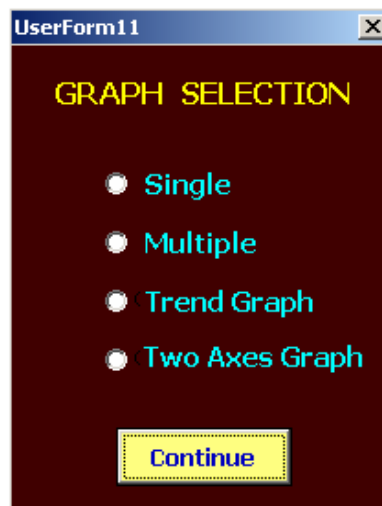
MTN time

The MTN selection box is used to synchronize the data across Pacific and Mountain Time zones. Data is stored in Arizona's time zone, and DataDoctor runs in Nevada's time zone. For half of the year, when Nevada goes on Pacific daylight savings time and Arizona goes on Mountain Time, there is a one hour time difference. The MTN box should be selected during that half of the year. Select the MTN box and save the DataDoctor application on your local drive to save the MTN setting. To restore the standard time settings six month later, select the MTN box to toggle the check mark off and save the application again.

3.3 After the Query – Creating Graphs and Transferring Data to HDB

3.3.1 Graph

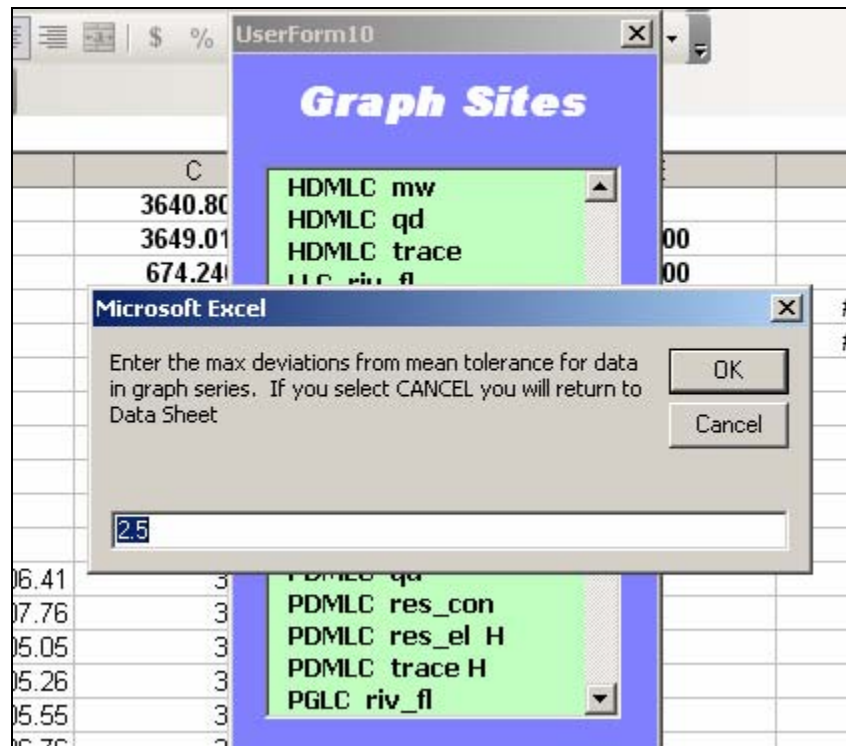
There are four graphing options in DataDoctor. When Graph is selected, a dialogue box pops up asking you to specify which type of graph you want to see:



Screenshot of Graph Menu Box

- Single - For single series graphs, you can select one site_datatype from those currently in the data sheet tab, as well as a color scheme for your graph – Blue, Red or Green. A line graph will be created with the parameters specified.
- Multiple - For multiple series graphs, you can select up to five site_datatypes from those currently in the data sheet tab. The scale of the y-axis ranges from the smallest data point value of the collective set of all parameters data points to the largest, so be careful to select variables that have comparable scales.
- Trend Graph – The trend graph works much like the multiple series graph, however the scale is standardized to make it easier for users to see how parameters change over time with respect to each other. The scale of the y-axis runs from -1 to 1. A box asking you to place a value for "Standard Deviations" appears when you run this graph. Usually 2.5 is appropriate for most graphs that do not feature chaotic data patterns, but for series data that fluctuate wildly, you can enter a lower number of standard deviations

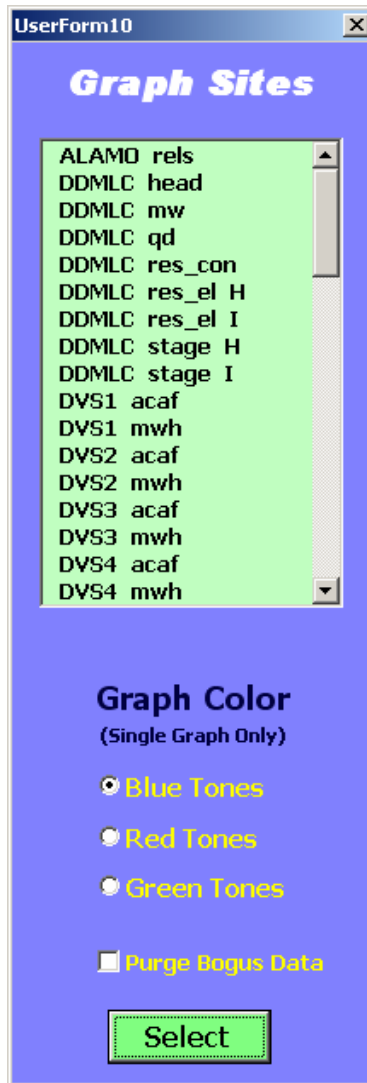
to help reduce the scaling effect of extreme outliers. This will enhance the visual impact of the graph, so the smaller values will not appear flattened.



Screenshot of Trend Graph Menu Boxes

- Two Axes Graph – Select the two axes graph to see two parameters on the same graph with two y-axes along the right and left sides. This option is appropriate when you want to see two data series with different scales, such as flow (cfs) and elevation (ft).

After selecting a type of graph from the menu, a userform labeled “Graph Sites” pops up. Near the bottom of the box is a checkbox labeled “Purge Bogus Data” (See screen shot below). If the “Purge Bogus Data” box is selected, the column that serves as source data for the graph (from the Data Sheet) will be purged of all data values that are associated with a blue cell. Eliminating blue cell values, which are usually out of limit or suspect in some way, will eliminate the y-axis scaling problems that may occur because of values that are far larger or smaller than the range created by normal data.



Screenshot of Site-Selection Dialogue Box, with Purge Bogus Data Option

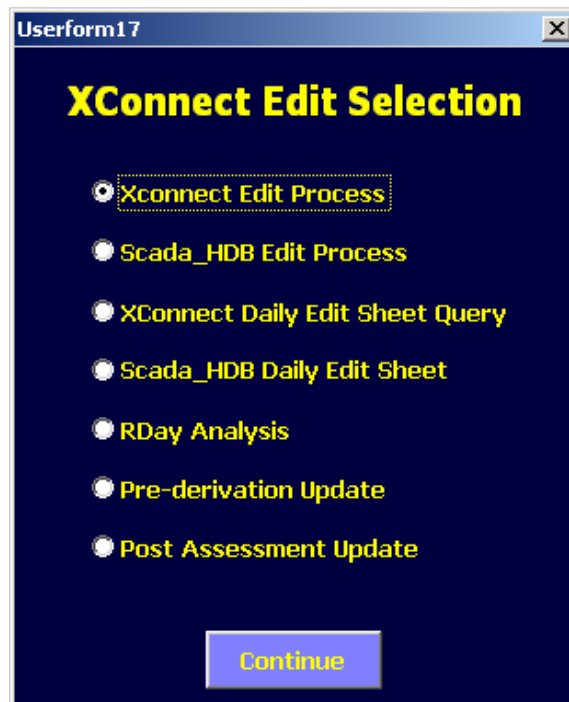
3.3.2 Transfer to HDB

Select Transfer to HDB after editing data in the data sheet tab. When Transfer to HDB is selected, a dialogue box pops up titled Continue Transfer to HDB. If you select 'Yes', DataDoctor runs SQL code to update the r_base table with any new or edited records. Upon completion, a dialogue box pops up informing you that the HDB update is complete. See HDB2 update Section 9 for more details on how DataDoctor uploads data to HDB2.

3.4 XConnect Edit

The XConnect Edit option is used in conjunction with one of several quality control application files (XConnect Daily Edit Records.xls, SCADA_HDB Daily Edit

Records.xls, or XConnect RDay Analysis Sheet.xls) to automate data tracking functions. The XConnect Edit function identifies missing or invalid data observations and updates records on the appropriate quality control application file. In addition, when the XConnect Edit Process option is selected, a list of data parameters that have missing or invalid data is created on the XConnect Edit worksheet tab. One row is created for each parameter-date combination with missing or invalid data. The hours of day that may contain bad data are listed across the columns. This list serves as a reference for manual data retrieval from the Sutron XConnect system. Additional information about the various XConnect Editing processes can be found in separate documentation (Filename TBD).



Screenshot for XConnect Edit Menu Box

3.5 XConnect Shift Coefficient Scan

The XConnect Shift Coefficient option provides data editors with a list of sites/parameters that have had recent shift and/or coefficient changes. The list of recent changes can be used by editors to determine if a shift or coefficient currently extant in the XConnect Desktop needs to be changed to the shift or coefficient value in effect during the time period to be polled. Failure to change shifts or coefficients from the current values to those in effect during the time period to be polled will result in bogus flow calculations for the site queried.

When a user clicks on the XConnect Shift Coefficient Scan option on the Main Menu, a message box pops up that asks if the DataDoctor Shift Register has been updated and if

the Shift Coefficient Scan.xls and Rating Table Coefficient and Site Flow Worksheet.xls files are both closed. If the user selects YES, a userform called XConnect Shift Coefficient Scan will pop up. The XConnect Shift Coefficient Scan form features two menu list boxes and a text box. One of the two menu list boxes is labeled Poll Groups and the other XConnect Sites. The Poll Groups listing offers a list of selections that includes XConnect Desktop poll groups such as HDMLC, BLKPT, and MTM. Each of these poll groups represent the various individual sites in the telemetry system that collectively make up the group. Only one poll group can be selected for a shift/coefficient scan. The XConnect Sites listing, on the other hand, offers a list of the individual sites available for polling via XConnect. Multiple sites can be selected from this list for each shift/coefficient scan. Users may select from only one list or the other. An error message will result if an attempt is made to select items from both the Poll Groups and the XConnect Shift menus.

The Cutoff Date textbox located toward the bottom center of the XConnect Shift Coefficient Scan userform contains a date that is used by the program code to determine how far back in time to evaluate XConnect Desktop shift and coefficient change occurrences. The automatically generated date that pops up initially with the userform, however, is not necessarily the cutoff date that should be utilized to do the scan. The user should determine the time period to be queried and then edit the cutoff date to reflect a time that is less than or equal to the beginning date of the query to be initiated.

After menu selections have been made and the cutoff date has been entered the user should click on the Continue command button at the bottom of the form. The program will then search for XConnect Desktop shifts and coefficients extant and in effect at various times back to the time of the cutoff date. Each qualifying shift and coefficient found by the program will then be placed and listed in a new spread sheet file called Shift Coefficient Scan.xls.

The new Shift Coefficient Scan spreadsheet will automatically appear on the screen when the program is finished creating it. The spreadsheet consists of two tabs or sheets. One sheet is called Shifts and the other is called Coefficients. The Shifts sheet, as the name implies, features a chronological listing of all the shifts that changed during the query period, the value of the shift, and the beginning date each shift was in effect. The Coefficients sheet features a chronological listing of all the coefficients that changed during the query period, the value of the coefficients, the beginning date each coefficient was in effect, and any comments that pertain to these coefficients. If the program search for either shifts or coefficients finds that there were no shifts or coefficients to report, a message indicating such will appear on the relevant sheet.

After the XConnect manual query is completed and/or there is no further need for the information of the Shift Coefficient Scan spreadsheet, the user can simply close the file without saving it.

4. RHourInstant Info and RDay Info Tabs

Both the RHourInstant Info and RDay Info tabs contain the same columns and serve the same purposes. They are the reference tables used by various macros in the DataDoctor application to perform queries, display formatting, and run quality checks. The tables contain the following column headings:

- Site – This is the site abbreviation. There can be multiple row entries with the same site.
- Parameter – This is an abbreviation for the attribute being measured
- Site-id – This is the HDB site_datatype_id. This field is necessary in order to perform queries and download fresh data from HDB to DataDoctor and upload new and revised data from DataDoctor to HDB.
- Analysis Order – This is the column order, from left to right, in which the data will appear in the data sheet tab whenever RHourInstant or RDay.
- Quick Look Order – This is a legacy column, and is no longer used.
- SQL Edit Route – This value tells DataDoctor where to query the data on HDB, and how to populate r_base fields when uploading data. On the RHourInstant table, a value of 1 means that the data should be retrieved from the r_hour table, and the interval, start_date_time, and end_date_time fields should correspond to hourly values. A value of 2 means that the data should be retrieved from the r_instant table and r_base fields listed previously should correspond to instantaneous values. On the RDay Info sheet, a value of 3 indicates that the data should be retrieved from the r_day table, but “day” is not the lowest interval in which the data is loaded into HDB2. In other words, the value is derived from hourly or instantaneous data, so any daily data loaded into r_base for this site_datatype needs to include an overwrite_flag of “O”. A value of 4 means that the data should be retrieved from the r_day table and “day” is the lowest interval in which the data is loaded into HDB2.
- Fmt plcs – This is a formatting specification, indicating the number of decimal places the retrieved data will display regardless of the format in which it is stored in HDB.
- Omit – Any row containing the number 1 under the Omit column will not be included in the analysis when Rhour-Rinstant Check or RDay Check (whichever is applicable) is selected from the data sheet menu options. The Omit column will either contain the number 1 or it will be null (indicating that DataDoctor should include the site_datatype in the analysis).
- Check Code – Contains the number 1 if the parameter is a shift (as in from the Shift Register), which indicates that DataDoctor should compare the HDB value with the value on the Shift Register tab. Otherwise contains the number 2, which indicates that regular quality checks (number of repeats, rate of change, etc) should be performed.
- Low Limit – This is the minimum value expected for the site_datatype.
- Hi Limit – This is the maximum value expected for the site_datatype.

- Repeat Limit – This is the maximum number of consecutive repeating values that may reasonably be expected to occur.
- ROC Limit – This is the maximum rate of change expected from a data value to the next consecutive value in time.
- Graph Units – This is the unit type that should appear as the y-axis label for the time series line graph for the datatype.
- Group Number – This value tells DataDoctor how to populate r_base fields when transferring data values to HDB2. In particular, each distinct group number corresponds to a group of values for the fields: agen_id, collection_system_id, method_id, and computation_id.
- Agency – This cell holds the value of the agen_id field for the particular site_datatype. Agen_id refers to the agency responsible for the original data record, or for providing updated values.
- Collection System Agency – This cell holds the value of the collection_system_id field for the particular site_datatype. Collection_system_id refers to the system that collected the data value.
- Method - This cell holds the value of the method_id field for the particular site_datatype. Method_id refers to the particular method, if any, that was applied outside of HDB to generate the data value.
- Computation ID - This cell holds the value of the computation_id field for the particular site_datatype. Computation_id refers to the particular computation, if any, that was applied outside of HDB to generate the data value.

Content Lookup Button:

On the top right corner of the RHourInstant Info tab is a button labeled Content Lookup. Select this button whenever you type in a reservoir elevation or forebay value in the high limit or low limit column for Hoover, Davis, or Parker Dam. The corresponding reservoir content for each of these facilities will be placed in the appropriate high or low limit cell on the sheet. This allows the user to quickly change a high or low elevation limit for any of the major dams without having to find a content book to look up the corresponding content values. The changes made to the RHourInstant Info sheet are also automatically copied to the RDay Info sheet to ensure congruity between the hourly and the daily parameters.

5. Shift Register Tab

The Blythe office River Operations Group takes periodic measurements of river (or canal) elevation and flow. Over time they have created a rating table, which maps an elevation measurement to a flow value. New measurements are taken every 3-4 weeks, at which time a shift, or deviation from the established rating table, may be noted.

For example, suppose the rating table maps an elevation of 302.0 ft to a flow of 4000 cfs, and an elevation of 302.6 ft to a flow of 4050 cfs. Now, suppose a new elevation measurement of 302.0 ft is found to have a flow of 4050 cfs. We see a shift from the

rating table of 0.6ft, because the 302.0 ft elevation now corresponds to the flow of a 302.6 ft elevation.

The date and time of the shift measurements are recorded and called in to the Boulder City office River Operations Groups, where they are entered into a Shift Register table, located on the shared network drive. Each user of DataDoctor should also store a copy of the shift register in their D:\DataDoctor\ directories. The Shift Register tab in DataDoctor consists of a copy of the Shift Register table.

The shift values in the table should match the values recorded into HDB2 from the telemetry system. The Shift Register tab is used for quality checking HDB2 shift values. If the HDB2 value for a given time interval does not match the value in the Shift Register table, the cell appears highlighted in yellow.

6. QuickLook Tab

The DataDoctor application allows users to create multiple QuickLook selections. When users select QuickLook from the main MENU, a dialogue box presents a list of different QuickLook parameter retrieval options. Users can select from one of the default QuickLook choices that come standard with DataDoctor, or they can create and customize QuickLook parameter options.

The QuickLook worksheet tab displays the same Sites, Parameters, and Site_Id columns (Columns A, B, and C) that are on the RhourInstant Info worksheet. Along row 4 of the worksheet beginning at column D you will see the names of the menu items that pop up when you select "QuickLook" on the main menu.

Use the following procedure to customize your personal QuickLook options:

Open your personal DataDoctorQuickLook.xls file. (For BCOO employees, that is the DataDoctorQuickLook.xls file in your D:\DataDoctor directory.) To make a new customized menu item click on the first cell on row 4 to the right of already established menu items and type in the name of the new QuickLook menu item you want to create. Then go down the column, looking at the sites and parameters in columns A and B. Type a 1 for the first parameter you want to appear on the data sheet screen during data retrieval, a 2 for the next, and so forth. The parameters will appear on the data sheet in the numerical order you specify in the column under your new menu item name. You can also add, remove, or change the order of parameters on the default QuickLook options, and you can delete any of the default options altogether.

When you have made all your changes, save the file. Then open your DataDoctor application file and update the QuickLook tab by selecting "Updates and Scada_HDB Edit" from the main MENU. Don't forget to save the DataDoctor application file to your

personal drive after updating your QuickLook tab so the new options will be available the next time you open the file.

If an updated version of DataDoctor is released with new site_datatypes that were not in use in the previous version, you can still update the QuickLook selections on the updated DataDoctor version by selecting “Updates and Scada_HDB Edit” from the main MENU. DataDoctor identifies the QuickLook columns in which each site_datatype should be listed by using the site_datatype_id in Column C of the personal QuickLook file. If a site_datatype_id does not exist in the personal QuickLook file, it will not appear in any QuickLook options on the new DataDoctor application. Users can add new site_datatypes to QuickLook options by using the procedure described above.

Note: QuickLook retrievals are only used for hourly and instant table retrievals. No QuickLook options currently exist for daily, monthly, or yearly data.

7. Avm Sheet Tab

7.1 Avm Sheet Reference Table

The Avm Sheet tab serves a similar purpose as the RHourInstant Info and RDay Info tabs. It is the reference table used by various macros in the DataDoctor application to perform queries and display formatting for AVM data. It contains some, but not all of the columns found in the other two reference sheets. The table contains the following columns:

- Site
- Parameter
- Site-id
- Analysis Order
- Quick Look Order
- SQL Table Route
- Check Code – Note: This column is null for all rows because quality checks are not performed the same way as they are for non-AVM data. See Section 6b on quality checks for AVM data below.
- Fmt plcs
- Omit
- Group Number
- Agency
- Collection System
- Method
- Computation ID

See Section 3 - RHourInstant Info and RDay Info Tabs for detailed descriptions of the column definitions.

7.2 Quality Checks for AVM data

The snapshot below displays a selection of the DataDoctor data sheet immediately following a fresh Avm Check:

	A	B	C	D	E	F	G	H	I	J	K	L	
1	MENU	RED	DVS1	DYS1	DVS2	DYS2	DVS3	DYS3	DVS4	DYS4	DVS5	DYS5	DDMLC
2			mwh	acaf	mwh	acaf	mwh	acaf	mwh	acaf	mwh	acaf	qd
3	Source = Interval												
4	REV	Repeat (long)			add>		add>		total >	0			
5		Repeat (sets)			add>		add>		minus>				
6	Missing Data Errors				add>		add>		minus>				
7	Outside Limits Errors				add>		add>		minus>				
8	Rate of Change Errors				add>		add>		minus>				
9	Max ROC Error				total >	0				0			
10	ROC Isolation			PASTE	avg >	#####							
11	End Time Values												
12	7/2/2006		-1.32	5	-1.40	0	47.73	376	48.87	386	39.32	364	13685
13	7/2/2006 1:00		-1.32	6	-1.40	0	47.98	376	48.61	381	-1.38	4	9281
14	7/2/2006 2:00		-1.32	5	-1.40	0	47.48	370	48.82	380	-1.41	3	9172

Select a range of cells and click on the RED button to change the font color of the data values to red. This aids the user by making ranges of cells with errors and/or edited values easier to identify.

Select a range of no more than five cells in a column and click on the PASTE button to copy and paste the selected data values to the column next to the left-most **add>** cells (under the column labeled DVS3 mwh). Directly below the fifth add cell will appear the sum of all values next to an **add>** cell. A total of ten values may be added; however, the sixth through tenth values must be manually copied and pasted to the column next to the right-most **add>** cells (under the column labeled DVS4 mwh). Below the total will appear the arithmetic mean of the data values.

The sum value will also appear to the right next to the **total>** cell (under the column labeled DVS5 mwh). Manually enter values below the total, next to the **minus>** cells to have that value subtracted from the total. The remainder appears in blue in the fifth cell down from the total. The remainder cell is cumulative; that is, if one value is entered next to a **minus>** cell, that one value will be subtracted from the total. If two, three, or four values are entered, all values will be subtracted from the total. No more than four values can be subtracted from the total to appear in the remainder cell.

A second set of PASTE, **add>**, **minus>**, etc cells appears to the right of the snapshot shown above for use with data from Parker Dam.

8. Quality Checks

Immediately after a query is completed a userform will pop up and ask the user if a data check of the queried data is desired. If the user decides to run a data check at that point, each data cell on the Data Sheet will be evaluated for a variety of data violations, as described below. If the user decides to forego the data quality check, the user is returned to the Data Sheet. Additionally, there is a Data Check option on the Main Menu. This option allows users to check data at their convenience without running a new query.

There are six check boxes on the Check Data userform; the user can select all or none of them:

- Check Repeat Values
- Check Limit Errors
- Check ROC Errors
- Purge Repeat Values
- Purge Limit Errors
- Purge Blue Cells

If selected, DataDoctor will perform basic quality checks as specified using the entries in the RHourInstant Info or RDay Info tables.

The data sheet tab displays the results of up to seven quality checks, depending on which checks were selected by the user:

1. Repeat (long) – If there is one set of consecutive rows containing identical values, this cell holds the number of consecutive repeating values appearing in the column. If there is more than one set of consecutive rows containing identical values, the cell holds the maximum number of repeating values; that is, the number of repeating values for the longest set. It is only populated if there are two or more repeating values in consecutive rows. The cell appears highlighted in orange, as do the cells in the column that contain repeated values.
2. Repeat (sets) – This cell holds the number of sets of consecutive repeating values that meet or exceed the Repeat Limit as specified in the RHourInstant Info or RDay Info tabs. Example1: Suppose the Repeat Limit is 1 and there are 10 data points – 10, 11, 11, 9, 11, 8, 8, 8, 7, 5, 8. There is one set of repeating values of 11 at rows 2 and 3, a second set of repeating values of 8 at rows 6, 7, and 8. Both sets of repeating values meets the Repeat Limit; therefore, the Repeat (sets) field will contain the number 2. Example2: Suppose the Repeat Limit is 6 and there are 10 data points – 2, 2, 2, 2, 2, 2, 2, 3, 2, 2. There is one set of repeating values of 2 that exceeds the limit at rows 1 through 7. The second set of repeating values, also of 2 at rows 9 and 10, does not meet or exceed the Repeat Limit. Therefore the Repeat (sets) field will contain the number 1.

3. **Missing Data Errors** – This value indicates the number of missing values for the `site_datatype` column. The cell appears highlighted in blue, as do the cells containing the missing values.
4. **Outside Limit Errors** – This value indicates the number of values that are outside the range of expected values for the particular `site_datatype`. The `RHourInstant` Info and `RDay` Info tabs both have columns labeled **Low Limit** and **Hi Limit**. If the value is lower than the entry in the **Low Limit** column or higher than the value in the **Hi Limit** column, the cell containing the value will appear highlighted in yellow. The **Outside Limit Errors** cell also appears highlighted in yellow if there are any values in the column that do not meet expected range criteria. Note: the **Outside Limit Errors** cell does not indicate whether the data values have exceeded the high limit criterion or the low limit criterion.
5. **Rate of Change Errors** – This value indicates the number of values for which the difference between it and the previous valid value (aka Rate of Change, ROC) is greater than the ROC Limit specified in the `RHourInstant` Info or `RDay` Info tab for the particular `site_datatype`. If there are any ROC errors, the cell appears highlighted in red. If there is only one data value with high ROC error, the cell containing that value is highlighted in green. If more than one value has an ROC error, the value with the highest ROC error appears highlighted in green and the other cells with the error appear highlighted in red.
6. **Max ROC Error** – DataDoctor finds the magnitude of the difference between each consecutive pair of values, and populates the **Max ROC Error** cell with the largest value. For example, suppose there are three rows populated with the values 500, 1000, and 2000 from top to bottom. DataDoctor subtracts the first value from the second and the second value from the third, and takes the absolute value of these differences to get ROCs of 500 and 1000, respectively. The maximum ROC is 1000, which gets populated into the **Max ROC Error** cell. Now, suppose there are three rows populated with the values 2000, 1000, and 500 from top to bottom. DataDoctor still subtracts the first value from the second and the second value from the third and takes the absolute value of these differences to get ROCs of 1000 and 500 respectively (not -1000 and -500). Again, the maximum ROC is 1000, so this value gets populated into the **Max ROC Error** cell.
7. **ROC Isolation** – This cell contains the number of rows between the two data values comprising the ROC error. If the data values are in consecutive rows, then the **ROC Isolation** row will contain the number 0. If there are multiple rows, for example, there is a valid data point followed by two invalid data points – missing or outside limits – followed by a valid data point that exceeds the ROC limit, then the **ROC Isolation** field contains the number of rows between the two valid points, in this case 2. If there are multiple ROC errors in the column, the **ROC Isolation** field contains the number of rows between values for the maximum ROC error.

Purging Cells

If Purge Repeat Values is selected, DataDoctor will delete all the values in cells highlighted in orange (recall, orange represents a value identical to that of the previous hour). The background will remain orange, although the contents will be empty. Users must select Check Repeat Values with Purge Repeat Values in order for the purge to function correctly.

If Purge Limit Errors is selected, DataDoctor will delete all the values in the cells highlighted in yellow (recall, yellow represents a value outside of operational limits defined by the user). The background of the cell will remain yellow, although the contents will be empty. Users must select Check Limits Errors with Purge Limit Errors in order for the purge to function correctly.

If Purge Blue Cells is selected, DataDoctor will delete the values in all cells colored blue, representing a missing or bogus value that did not get derived to the interval tables. Values will only appear in blue cells if the user specifies that DataDoctor should query from r_base. When the Purge Blue Cells function runs, the background remains blue although the contents will be empty. Thus the data sheet will appear as though the user had specified that DataDoctor should only query from the interval tables. The Purge Blue Cells option functions independently of any other data checks; therefore users may select this option without selecting any other data checks.

For all purge functions, values are deleted from the spreadsheet data view only. The records in HDB are not affected by DataDoctor purge functions.

9. Updating Changes to HDB2

DataDoctor users can upload any changes to the data that they have made manually by selecting the Transfer to HDB option from the data sheet main MENU. When you select the Transfer to HDB option, a dialogue box pops up asking you to confirm that you want to continue the data transfer. Near the bottom of the box is a checkbox labeled “Purge Bad Time Stamps.” If this box is checked the program searches through the record set that gets loaded to HDB and removes records with time values that do not correspond to the top of the hour.

Updates to HDB can only be done when the most recent data retrieval came from r_base, r_instant, r_hour, or r_day. Monthly and yearly data cannot be directly transferred to HDB using the DataDoctor application.

As a security and integrity upgrade for HDB2, DataDoctor only has permission to enter data into the r_base table. The HDB2 Derivation Application is the only application that can update time series tables such as r_hour and r_day.

When Transfer to HDB is selected, DataDoctor searches the data sheet for any cell with a red background. Whenever it finds such a cell, it looks up the site_datatype_id for the column and uses information in the RHourInstant Info or RDay Info tab to create a

complete record that can be inserted into the r_base table. Data values that do not appear with a red background will be passed by and will not get placed into the record set that gets uploaded to HDB2. Once the transfer to HDB2 is complete, all the values that were updated change from red backgrounds to purple backgrounds. If the value in the spreadsheet cell is identical to the value currently existing in the r_base table, the record will not be updated. To indicate the record was not updated, the background color will not change to purple, but will remain red.

After transferring new data to HDB2, the user can continue to edit values in the data sheet. This is the case for data values that were left unchanged in the transfer as well as values that were already edited one or more times. The background color of an altered cell will continue to change to red, whether the current background is purple, white, blue, or any other color. There is no limit to the number of times a user can make changes and update data to HDB2 in a single DataDoctor session.

Purge Bad Time Stamps

Sometimes data with time stamps that do not correspond to top of the hour find their way into HDB. When this happens, the program code may skip one or more hours of data that should be transferred to HDB. When this happens, DataDoctor immediately stops transferring data from Excel to HDB.

If the “Purge Bad Time Stamps” box is checked, DataDoctor will search through the records in HDB and ignore records with time stamps that do not correspond to top of the hour before replacing HDB values with revised values from DataDoctor. For the sake of efficiency, users should first attempt to transfer data to HDB with the “Purge Bad Time Stamps” box unchecked. If the user leaves the purge option unchecked, DataDoctor does not check the record set and the transfer is processed as described in the section above. If the transfer process gets stuck, the user should exit the procedure and try the transfer again, this time using the Purge Bad Time Stamps selection. If the transfer fails again there is some other problem that must be addressed. Contact Joe Donnelly (702-293-8319) or Rich Hedrich (702-293-8356) in this case.

10. Spreadsheet Security

In response to a need to improve security for the DataDoctor and its affiliated spreadsheets, security features were added to version 3.3. The DataDoctor spreadsheets protected with new security code include the DataDoctor RhourInstant Rday.xls (control file), XConnect Daily Edit Records.xls, and XConnect RDay Analysis Sheet.xls.

The DataDoctor and affiliated files are protected from unauthorized use by requiring users to provide a DSN, UID, and PWD entry prior to any attempts to query or transfer data to HDB. At DataDoctor start up a userform called Data Base Login & Admin Path will pop up on the screen. This userform contains several text boxes that serve as the entry points for the DSN, UID, PWD, Administration path, and the backup path.

The DSN entry is the name of the default, or most recently accessed data base, and is normally already available in the DSN textbox when the userform pops up. The only time this entry should be changed is if the user wants to do work with a different data base. If a new data base name is typed into the textbox this name will be remembered by the DataDoctor as the default database and pop up in the DSN textbox the next time the login userform is summoned.

The UID entry is the user's identification name. This entry is also already available in the UID textbox when the userform pops up. If this name is changed the DataDoctor will remember it and place it in the UID textbox as the default entry the next time the login userform is summoned.

The PWD textbox does not contain an entry at initial startup. The user is required to provide this information. When a password is typed into the text box cryptic “ * “ characters will appear instead of letter, numbers, etc. to maintain security. If no password is provided and the user clicks the Continue command button a message will pop up on the screen informing the user that no password was provided...try again. If a bogus password is typed into the text box the program will accept it and continue on as though nothing is wrong. However, if an attempt to transfer or query data to or from HDB is attempted the user will not be successful. A message will pop up on the screen informing the user that a bad password, login information, or sdid was encountered by the program code...and to login again with the correct login information.

The Admin Path text box is the name of the path and directory where the DataDoctor file and the DataDoctor control files reside. This path is the default, most recently used path and is present in the text box when the Login userform pops up. As with the other text boxes on the userform the DataDoctor will remember any changes to the path you happen to make. Some of the tasks the Admin Path is required for include updating QuickLook, the Shift Register, and the Rhour RDay Info sheets.

The Backup Path (hard drive path) is the name of the path and directory on the user's hard drive that harbors backup and/or customized Shift Register and QuickLook files.

The DataDoctor takes information off of these customized hard drive files when the user wants to update his or her personal version of the DataDoctor with a customized Shift Register, or a QuickLook that contains all of the latest QuickLook control file sites/parameters, limits, rates of change etc. as well as the customized listings available on his or her customized hard drive version of the QuickLook file.

If needed, the Login form that appears at initial file startup can also be accessed by clicking on the MENU command button on the Data Sheet and selecting the Data Base Login option. After clicking Continue on the Main Menu the Login userform will pop up just as it does at startup. If the password originally placed in the PWD textbox is still available to the program code the box will be filled with “ * ” characters. There is no need to re-type a password unless the password originally typed into the text box was incorrect. If the box is blank the user should re-type the password into the text box.

Access to the Login userforms associated with DataDoctor affiliated spread sheets such as DataDoctor RhourInstant Rday.xls, XConnect Daily Edit Records.xls, and XConnect RDay Analysis Sheet.xls are accessed via a LOGIN command button located on one the spread sheet sheets/tabs of each file. On the DataDoctor RhourInstant Rday.xls file, the Login command button is located on the “Data Sheet” tab. On the XConnect Daily Edit Records.xls file the button is located on the “XConnect Update Record” tab. On the XConnect RDay Analysis Sheet.xls file the button is located on the “Summary Sheet” tab.

As with the DataDoctor, Login access information for DataDoctor affiliated files is not required to work on the spread sheet, but rather to utilize command button automated procedures that require communication with HDB. Some of these automated procedures include querying and displaying HDB limits for data editor analysis, finding HDB site/parameter method id’s, agency id’s, computation id’s, and collection system id’s, and general data editing HDB data queries.

11. Glossary of Frequently Used Terms

Term	Definition
AVM	Acoustic Velocity Meter; a type of telemetry equipment that measures water flow through a pipe, penstock, or other waterway.
BCOO	Boulder Canyon Operations Office (located in Boulder City, NV)
HDB / HDB2	Hydrologic Database, Version 2; implemented at BCOO in 2004
ROC	Rate of change
SCADA_HDB	The SCADA_HDB file is a text file sent to the ib3r3watr UNIX server by the Hoover Dam SCADA system every hour. It contains hourly data for 47 parameters, mostly Hoover Dam data, but a few Davis Dam and Parker Dam parameters as well.
site_datatype	Unique combination of a data telemetry site and a specific data type. Examples include Lake Mead Elevation, Lake Mead Content, Lake Havasu Elevation
site_datatype_id	Unique 4-digit identifier for every site_datatype in HDB
DSN	Database Source Name
UID	User Name for login to LCHDB
PWD	Password for LCHDB