



# STANDARD UPLOADING GUIDE

**Time Series and Sensor Data Publication on [data.cuahsi.org](http://data.cuahsi.org)**

**Version 2.0**

**Prepared by  
Liza Brazil, Community Support Specialist**

**February 2018**

## CONTENTS

Contents .....	2
Disclaimer .....	3
Introduction and Purpose of This Document .....	4
Workflow for Publishing Data .....	4
Create Publishing Account .....	4
Upload Data .....	5
Upload Metadata .....	5
Upload Data Values .....	8
View and Edit Metadata in the Database .....	8
Publish Your Data .....	9
Appendix A: The Observations Data Model (ODM) .....	9
Appendix B: Required and Optional Tables .....	11
Appendix C: Controlled Vocabularies in the ODM .....	11

## DISCLAIMER

*The CUAHSI Water Data Services is funded by the National Science Foundation to provide services to discover, retrieve, and publish water data. This material is based upon work supported by the National Science Foundation (NSF) under Grant No. 13-38606. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.*

## INTRODUCTION AND PURPOSE OF THIS DOCUMENT

The CUAHSI Hydrologic Information System (HIS) enables data uploading ([HydroServer](#)) and data discovery ([HydroClient](#)). This document describes how to create a publishing account, upload data, and request publication to make your data accessible to the public on HydroClient. Before proceeding with uploading your data, you must format it to fit with one of the two CUAHSI formatting options: The Standard Format Option or The Advanced Format Option. The Standard Format Option is a subset of the Advanced Format Option. Instructions on how to format your data and which format option to select are provided in the [CUAHSI Formatting Guide](#).

The HydroServer provides two uploading options that correspond to the two options for formatting data. The Standard Upload option involves uploading only the required tables: Variables, Sites, Sources, Methods, Quality Control Levels, and Data Values. Refer to the Advanced Uploading Guide for instructions on uploading the full metadata profile, including all required tables and any optional metadata tables, e.g. Samples, LabMethods, Categories, Qualifiers.

## WORKFLOW FOR PUBLISHING DATA

How to publish data in CUAHSI HIS:

1. Create a publishing account
2. Verify your account using the activation email
3. Format your data: Follow the Guide and fill in the Excel Templates ([Standard](#) or [Advanced](#))
4. Upload your data: Follow the Guide (Standard or [Advanced](#)) to complete the process
5. Request Publication: Click Request Publication on the HydroServer Homepage to make your data service accessible to the public on <http://data.cuahsi.org>
6. For all subsequent uploads, data is published every Saturday automatically unless requested otherwise

Questions? Contact CUAHSI support at [help@cuahsi.org](mailto:help@cuahsi.org)

## CREATE PUBLISHING ACCOUNT

You will need a Google Account to create a publishing account. The Google Account you use when creating a publishing account will be associated with the HydroServer. You may decide to create a shared Google Account for your organization or you may have several individual Google Accounts linked to one database.

To create a publishing account, go to <http://hydroserver.cuahsi.org/> and click Create New Publishing Account. Fill out the required account information as described below.

- Service Name: An abbreviated name for your data service. Example: NWISDV
- Service Title: The full name of your data service. This will be displayed to the public. Example: NWIS Daily Values
- Service Abstract: A detailed description of your data service and the data.
- Contact Name: Name of the person that the public should contact with questions about the data.
- Contact Email: Email of the person that the public should contact with questions about the data.

- Contact Phone: Phone number of the person that the public should contact with questions about the data.
- Organization: Name of the organization responsible for publishing the data.
- Organization Website: Website address of the organization responsible for publishing the data.

**Note:** CUAHSI Staff will use the information from your publishing account to create an [HIS Central Data Service](#) page. This will list you among all data service providers in the CUAHSI Catalog. CUAHSI Staff will provide you with credentials to edit this information at any time.

## UPLOAD DATA

Standard uploading is broken up into two processes: (1) upload metadata files (2) upload data values files. The process is identical for both metadata and data value uploading but it is important that you first upload metadata before moving on to data values.

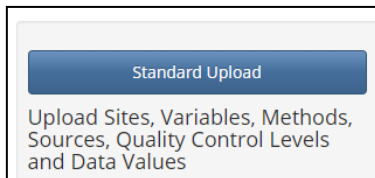
1. Browse to or drag and drop all metadata files as CSVs
2. Upload files to first validate that your metadata files have the correct and required headers from the CUAHSI Formatting Guide
3. Continue to insert valid metadata files and check for errors in the formatting
4. Fix up to 25 rows of errors in the browser and re-upload
5. Repeat for data values once associated metadata values are inserted

---

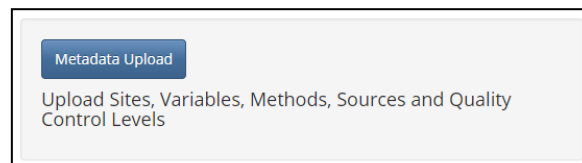
## UPLOAD METADATA

---

To get started, click on the Standard Upload option as shown below.



You are required to upload all metadata before uploading any associated data values. Click on the **Metadata Upload** button as shown below.



Drag and Drop all metadata files as shown below.

## Upload Metadata

The screenshot shows a 'Drag & Drop' area with five rows of file information. Each row contains a file name, a progress indicator (a small circle), and a 'Remove' button. The files are: Methods01.csv, QualityControlLevels.csv, Sites.csv, sources.csv, and variables.csv. Below the list are three buttons: '+ Add Files' (blue), 'Cancel' (orange), and 'Upload' (green).

Click **Upload** or click **Remove** if you have selected the wrong file. The HydroServer will load the files and then check to ensure that:

1. The files contain the correct headers required in the Standard metadata templates
2. The files are not missing any headers required in the Standard metadata templates

Files with errors will appear red and provide a description of the error. For example, below shows a Methods file that contains an error because it is missing the MethodDescription header that is required in the Methods table.

## Upload Metadata

The screenshot shows the same 'Drag & Drop' area, but now with validation results. The 'Methods01.csv' row is highlighted in red and has a red bar with the text 'Validates as: "Methods" with 1 missing column name(s)'. Below this bar is a text box containing 'Missing headers: MethodDescription'. The other four rows (QualityControlLevels.csv, Sites.csv, sources.csv, and variables.csv) are highlighted in green and have green bars with text indicating successful validation: 'Validates as: "Quality Control Levels" with 4 record(s)', 'Validates as: "Sites" with 215 record(s)', 'Validates as: "Sources" with 2 record(s)', and 'Validates as: "Variables" with 3 record(s)'. Below the list are four buttons: '+ Add Files' (blue), 'Cancel' (orange), 'Upload' (green), and 'Insert' (blue).

You are required to **Remove** all files that have errors before proceeding on. Click **Insert** only after all files have been loaded and contain no validation errors. Once all files have been loaded and validated with no errors, click **Insert**. Once file insertion is complete, click **Continue**. After clicking **Continue**, the HydroServer will check the contents of the files. The Upload Summary describes the number of records by table that are:

- **New:** Valid records ready to be committed to the database.
- **Rejected:** Records that could not be validated by the application.
- **Updated:** Records that have been updated because the uploaded record had a previously used key with other, changed metadata that are ready to be committed to the database.
- **Duplicate:** Records that are a complete match of records that have been previously committed to the database.

**Note:** The HydroServer Uploader reports on and does not insert duplicate metadata, however the HydroServer does not report on duplicate data values. Instead, all valid data values are inserted and duplicates are checked for and removed automatically on a weekly basis. This process helps to increase the speed of the upload.

An example of the Upload Summary table is shown below. Notice there are 3 records that have been rejected in the Sites table and 2 records in the methods table.




## Upload Summary

Table Name	Inserted	Updated	Duplicated*	Rejected
<b>Sites</b>	212	0	0	3
<b>Variables</b>	1	0	2	0
<b>Sources</b>	2	0	0	0
<b>Methods</b>	2	0	0	2
<b>QualityControlLevels</b>	0	3	1	0

[Finish](#)

Edit the Rejected records by clicking on the rejected column. A pop up window will appear with a red drop down arrow that allows you to edit the row. Click on the red arrow to view information on why the record was rejected. In this example, the LatLongDatumSRSName is empty. Click to select the cell under LatLongDatumSRSName, to enter the correct LatLongDatumSRSName as shown below. This example shows the correct LatLongDatumSRSName as WGS84.

Rejected Sites

SiteCode	SiteName	Latitude	Longitude	LatLongDatumSRSName	SiteType	Comments
 wq1	Fyrisån Flottsund	59.786938	17.659863	<input type="text" value="WGS84"/>	Stream	
<small>The value in column LatLongDatumSRSName can not be empty</small>						
 wq2	Fyrisån Klåstorp	59.886636	17.578593		Stream	
 %%%%	Sälvaån Ingvasta	60.01157	17.848522	WGS84	Stream	

Click on the remaining rows to review the error and edit the necessary cells. Once all rows are corrected, you may click **Re-Upload**. Alternately, you may decide to click **Download** to correct the errors in Excel before re-Uploading. Once the data has been re-uploaded, the Upload Summary should reflect the corrections and show the increased number of Inserted records. Click **Finish** to navigate back to the HydroServer Homepage and continue to upload Data Values.

**Note:** You may edit up to 25 records at a time. For files that have more than 25 records with errors, you are required to download the file and re-upload once corrected. The downloaded file will contain a column with errors describing why the record was rejected.

---

## UPLOAD DATA VALUES

---

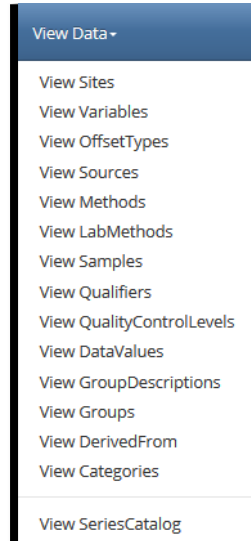
To upload the data values, follow the same steps as for uploading metadata. To get started, navigate back to the Standard Upload page and click on **Data Values Upload**.

---


## VIEW AND EDIT DATA IN THE DATABASE

---

To view data that have already been loaded into the database, click on View Data in the top menu, then select the table that you wish to view.



The resulting webpage will allow you to view, search, and sort through data that have been committed to the database.



**HydroServer**  
CUAHSI Open Source Hydrologic Data Tools

Home Contact Manage Templates Controlled Vocabularies Import CSV Data View Data Admin Hello lizabrazil505@gmail.com!

### Sites

Show 10 entries Search:

Site Code	Site Name	Latitude	Longitude	LatLongDatumRSName	Elevation_m	VerticalDatum	LocalX
105	Outer Harbor 105	42.326299	-70.948951	Unknown			
106	Outer Harbor 106	42.330614	-70.966117	Unknown			
117	Hingham Bay 117	42.291261	-70.916335	Unknown			
118	Outer Harbor 118	42.331629	-70.936248	Unknown			
124	Hingham Bay 124	42.275767	-70.899513	Unknown			
129	Winthrop Bay 129	42.356751	-70.98088	Unknown			
130	Winthrop Bay 130	42.365376	-70.991008	Unknown			
137	Mystic mouth 137	42.386934	-71.065337	Unknown			
138	Mid-inner harbor 138	42.361063	-71.040962	Unknown			
139	Quincy Bay 139	42.280086	-70.970237	Unknown			

Showing 1 to 10 of 46 entries First Previous 1 2 3 4 5 Next Last

Report a bug here! © 2016 - © CUAHSI Version 1.1.5

Editing the contents of a database requires a new upload. For Sites, Variables, Methods, Sources, and QualityControlLevels, uploading a record with a previously used code (SiteCode, VariableCode, MethodCode, SourceCode, and QCLLevelCode) will update the corresponding record in the database with



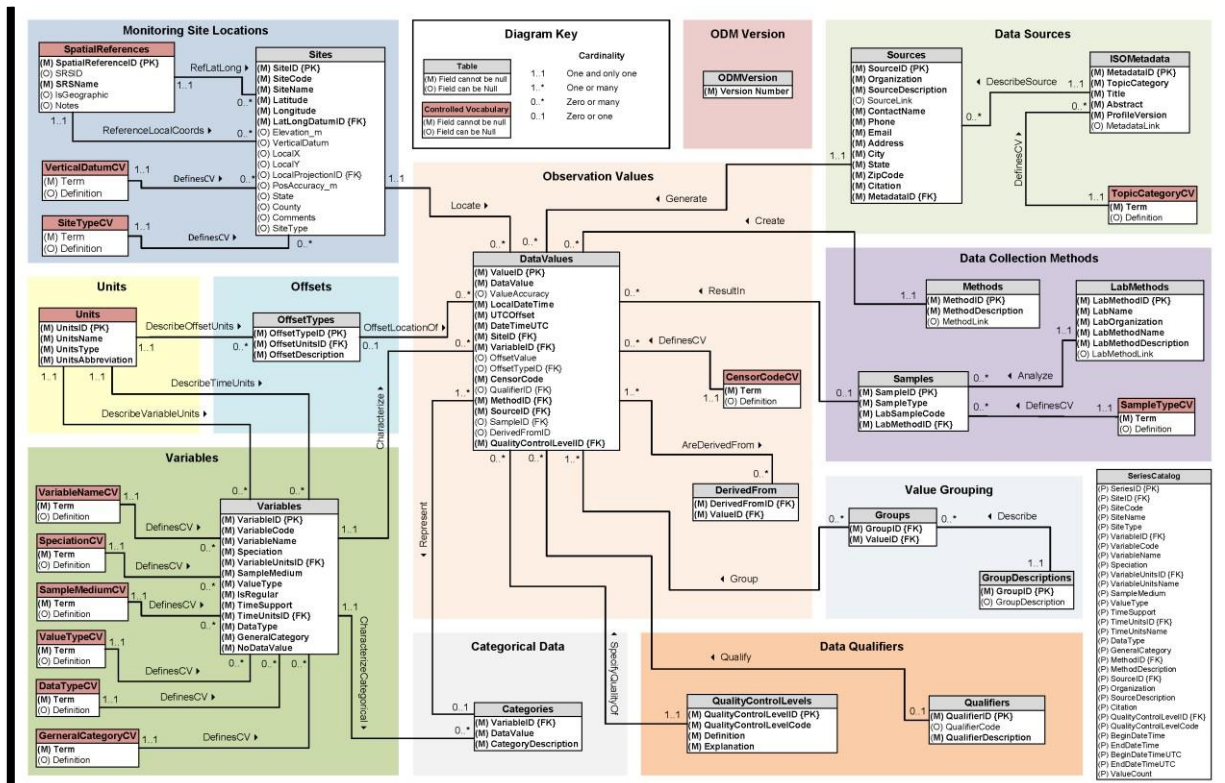
the information from the new upload. Such records will be presented in the validation step during upload under the *Updated* column. Edits to the Data Values must be performed by contacting [help@cuahsi.org](mailto:help@cuahsi.org).

## PUBLISH YOUR DATA

Once you are ready to publish your data to be discovered on <http://data.cuahsi.org/>, click **Request Publication** on the HydroServer homepage. A request will be sent to [help@cuahsi.org](mailto:help@cuahsi.org) to publish your data. Expect a response from [help@cuahsi.org](mailto:help@cuahsi.org) with a link to review your data on a private version of <http://data.cuahsi.org/>. Once you have reviewed and confirmed to [help@cuahsi.org](mailto:help@cuahsi.org), CUAHSI staff will publish your data and create an HIS Central Data Service page that contains the information you provided when creating a publishing account. CUAHSI Staff will provide you with login information to access your HIS Central Data Service page and make any edits to the information.

**Note:** Data services are harvested once a week. If a data publisher needs their data to be harvested prior to the next scheduled harvest they can send a request to [help@cuahsi.org](mailto:help@cuahsi.org).

## APPENDIX A: THE OBSERVATIONS DATA MODEL (ODM)



The Observations Data Model (ODM) is the information model employed by the CUAHSI HIS. The most common implementation of it is as relational database in Microsoft SQL. The core of the model comprises of a center table that contains the value of observations as well as foreign keys to ancillary tables. These tables provide metadata with the goal of unambiguous interpretation of the data values and include tables with information related to the:

- Location of the observation (*Sites*)
- Phenomenon being observed (*Variables*)
- Methods being employed (*Methods*)
- Sources of the data (*Sources*)
- Quality control techniques employed (*Data Qualifiers*)

The ODM is described in a peer-reviewed article in [Water Resources Research](#) and can be accessed [here](#).

## APPENDIX B: REQUIRED AND OPTIONAL TABLES

Table Name	Table Description
<i>Variables</i>	The Variables template lists the full descriptive information about what
<i>Methods</i>	The Methods template lists the methods used to collect the data and any
<i>Sites</i>	The Sites template provides information giving the spatial location at which
<i>Sources</i>	The Sources template lists the original sources of the data, providing
<i>Samples</i>	The Samples template gives information about physical samples analyzed in
<i>LabMethods</i>	The LabMethods template contains descriptions of the laboratory methods
<i>Quality Control</i>	The QualityControlLevels template contains the quality control levels that
<i>DataValues</i>	The DataValues template contains the actual data values and keys to
<i>Categories</i>	The Categories template defines the categories for categorical variables.
<i>DerivedFrom</i>	The DerivedFrom template contains the linkage between derived data
<i>GroupDescriptions</i>	The GroupDescriptions template lists the descriptions for each of the
<i>Groups</i>	The Groups template lists the groups of data values that have been created
<i>Qualifiers</i>	The Qualifiers template contains data qualifying comments that accompany

Required tables are orange and optional tables are green in the above.

## APPENDIX C: CONTROLLED VOCABULARIES IN THE ODM

Within the ODM, certain fields have been designed as controlled vocabularies, which are community driven and moderated by CUAHSI staff. Click here to visit the landing page for the [Master Controlled Vocabulary Registry for ODM 1.1](#).

The list of fields that are controlled vocabularies can be seen below. These fields are listed by template and hyperlinked to the appropriate web page to view, submit edits, or submit new terms:

### Variables

- [VariableName](#)
- [Speciation](#)
- [VariableUnitsName](#)
- [SampleMedium](#)
- [ValueType](#)
- [TimeUnitsName](#)
- [DataType](#)
- [GeneralCategory](#)

### Sites

- [LatLongSRSName](#)
- [LocalProjectionSRSName](#)

## Sources

- [TopicCategory](#)

## Samples

- [SampleType](#)

## DataValues

- [CensorCode](#)