Interoperable Watersheds Network

User manual for IWN Data Appliance

February 8, 2018

Contents

Figures

The mission of the United States Environmental Protection Agency (EPA) is to protect human health and the environment, including air, water and land. Understanding the extent of pollution in waters and identifying waters for protection has been based in part on water quality monitoring data collected and shared by parties (federal, state, tribal, and local) throughout the U.S. The purpose of EPA's Watersheds-based Monitoring and Interoperable Data Platforms project was to pilot the development of a centralized catalog for discovering sensor data. The site provides a single portal, where participants could register their sensors, and be included in data searches based on parameters collected or geographic attributes.

Installation

The easiest method for deploying the Interoperable Watersheds Network data appliance is by using Docker and docker-compose which are supported on Windows 10, Linux, and MacOS. If preferred however the individual components may also be installed individually. The below commands are given as examples only and are structured for linux-type environments. You may need to modify them to meet your specific needs.

Step 1: Download Data Appliance from GitHub Repository

Pull the latest code from the GitHub repository located at https://github.com/USEPA/Interoperable-Watersheds-Network-Data-Appliance, ex:

```
git clone https://github.com/USEPA/Interoperable-Watersheds-Network-Data-Appliance
```

This will retrieve a copy of the latest code from EPA's GitHub repository. It will create a folder in your home directory titled 'Interoperable-Watersheds-Network-Data-Appliance'. All the scripts and code are located in this directory. The Data Appliance has been designed to be deployable using Docker, which provides a container for the application along will all the required dependencies that the application needs to run. You can find out more about docker at: https://www.docker.com/.

Once the data appliance has been downloaded, you could go straight to step 4 to begin standing up your instance, however, this uses all the default settings. Although this may be fine for testing how the data appliance works, for production purposes, you will want to modify several settings before proceeding with initiating the instance.

Step 2: Security

The supplied images are for demonstration and testing purposes and are supplied without sufficient security in place to be used for production use. The user interfaces and services should be limited to secure networks and/or sufficient authorization implemented within the webserver by htpasswd or similar. A user with the url for the data appliance would be able to access the data appliance UI directly if you do not employ some user/password approach as described above.

The password for the database that stores the data ingested by the data appliance should also be changed to avoid someone being able to change and/or delete your data. The default admin password for the SOS interface is ***52northAdmin*** and should be changed. The database password is specified in db/scripts/1init.sql and should be changed.

If deploying to a production environment, ports 80, 8000 and 8080 should be open or a proxy should be implemented. It is not recommended or necessary to open port 5432 for the database.

Step 3: Customization

The default organization information and all lookup values seen in the user interface can and should be modified before installation by editing the sql scripts located in db/scripts. You should start by modifying the organization information to match how you want your organization to be represented in the services that are publishing your data. This can be modified by changing the 5org.sql script.

```
INSERT INTO sos.organizations (organization_id, parent_organization_id, name, url,
contact_name, contact_email, sos_url)VALUES('example', 'example', 'Example Organization',
'www.example.com', 'John Doe', 'jdoe@example.com', 'http://ingest_sos:8080/52n-sos-
webapp/service');
```

db/scripts/5org.sql – Change the highlighted values to match your organization's information. 'Ingest_sos:8080' must be changed to your server's url where the data appliance has been deployed.

The organization URL should be changed to reflect the deployed instance of the SOS if in a production environment. By default, it is pointing to the Docker container's instance (http://ingest_sos:8080/52n-sos-webapp/service). Take note of the organization_id that you use here as it will be needed to access your data appliance user interface.

Lookup values such as parameter names, units, etc., are contained in the 3data.sql script. A default set has been provided, but you can provide your own. These will only be added upon the initial composition of the docker instance. Adding domains after this time will need to be done by access the 52North PostgreSQL database directly and adding value to those tables.

Name	value
service.sosUrl	http://localhost/52n-sos-webapp/service
serviceProvider.site	www.example.com
serviceProvider.name	Example Company
serviceProvider.individualName	John Doe
serviceProvider.positionName	IT Specialist
serviceProvider.email	john.doe@example.com
serviceProvider.address	1313 Mockingbird Ln
serviceProvider.city	Beverly Hills
serviceProvider.state	CA
serviceProvider.country	United States

sos/configuration.csv - edit this file to provide your organization's information as you'd like it to be presented within the web service.

The SOS interface also contains variables that should reflect your organization (admin->settings) and can be modified after installation using the web interface. You can also edit these settings prior to installation by changing the configuration.csv file located in the sos folder.

The last configuration file that needs to be changed is located in web/scripts/config.js. You will need to update this file to reflect the url that will be used for the web services that the data appliance uses to ingest

data. By default, it is set to localhost. You will need to modify this to match the url of your environment. Only one serviceUrl should be active.

```
var config = {
    //*** local *** change localhost if deploying to a different environment
    serviceUrl : "http://localhost:8000/"
}
```

web/scripts/config.js - modify this file to match your environment.

If you choose to use a different database location (not the one that is created as part of the docker compose process), you must also change sos/datasource.properties and src/app/ingest/ingest.py to point to your database. This is not a required step if you choose to go with the default database that is created as part of the installation process.

Step 4: Compose the Data Appliance Using Docker

Using Docker (note: you may need to install docker first if it is not currently part of your environment):

```
cd Interoperable-Watersheds-Network-Data-Appliance
docker-compose up
```

When Docker is finished pulling and building the images/containers, the user interface will be available at http://localhost/ingest.html?orgId=example. The SOS interface will be available at http://localhost:8080/52n-sos-webapp. The REST services will be available at http://localhost:8000. Note that if you modified the configuration in step 3, then 'localhost' will be replaced with the URL that you specified. 'example' will be replaced with the organization_id that you specified.

The data stored in the database is persisted to a Docker volume and will remain if the container is stopped and or rebuilt unless the volume itself is removed via Docker commands. It is recommended for production use that the database be deployed to a suitable location with proper backup and recovery, etc.

Requirements for deploying individual components (Docker retrieves all of these for you, but if you choose to install the data appliance without Docker, this is what you will need):

PostgreSQL 9.6+

Tomcat 8+

Python 3+

Nginx or another suitable webserver

Interacting with the User Interface

Reviewing Organization Information

The User Interface will provide you information about your organization. You set this up in Step 3 above. This can be changed after installation in the SOS interface of 52North within the data appliance. This guide does not describe how to do that.

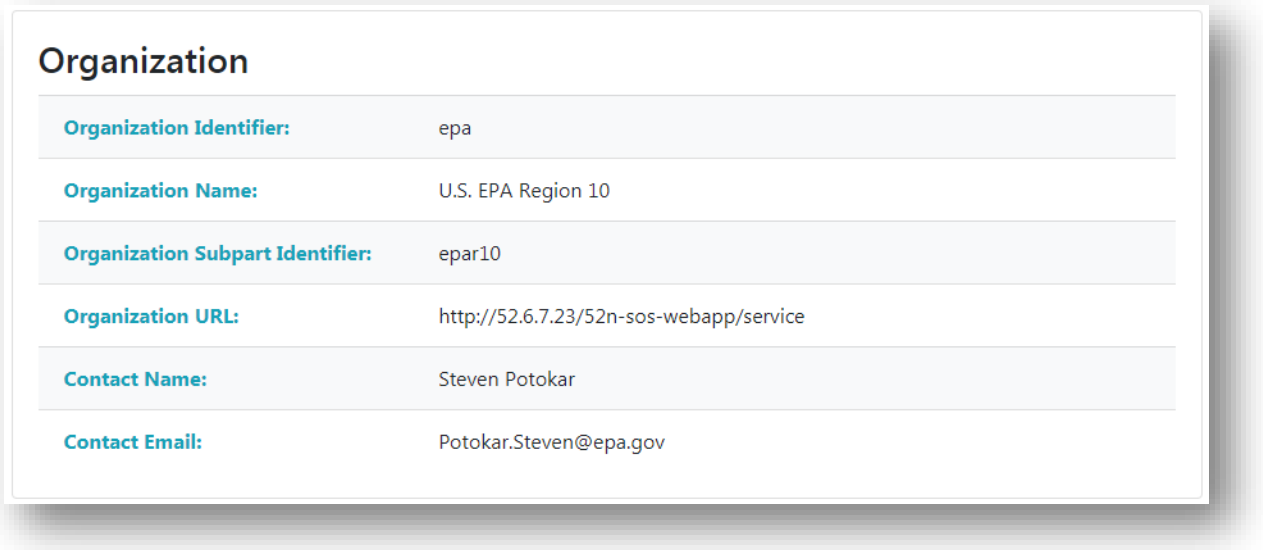

Figure 1. Organization data of registered user

Registering a Sensor

Sensors added by the current user are listed in the **Sensors** grid by ID and Sensor Name (Figure 2). Select a column header to sort ascending or descending by the selected column. Filter the sensor grid by entering text in the Search field. The grid automatically updates to display filter results. Clear the Search field to reset the Sensor grid.

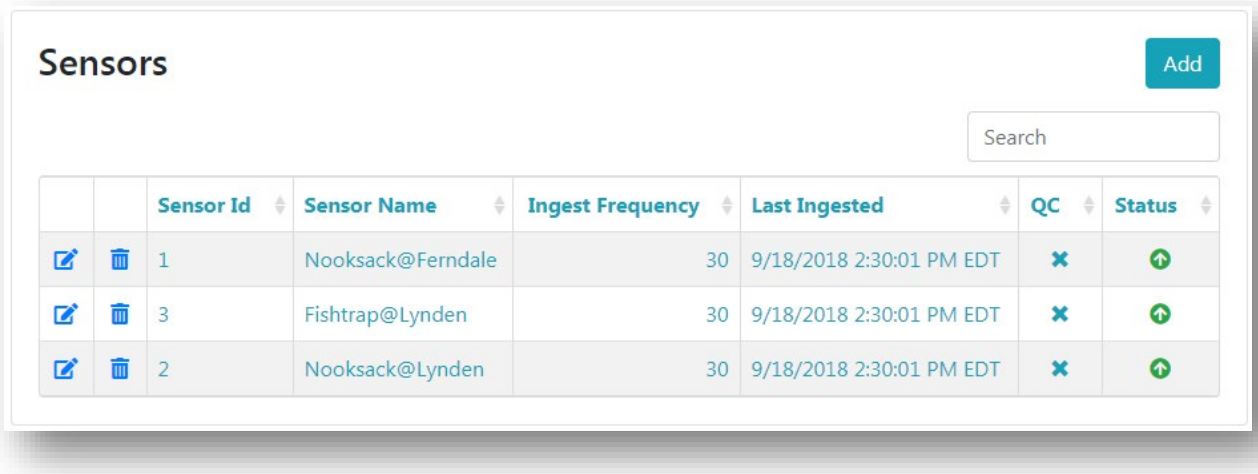

Figure 2. Sensors grid

The Ingest Frequency column displays how often data is retrieved from the sensor in minutes. The recommended ingest frequency is 30 minutes. The QC (quality control) column displays any quality control applied to a sensor to QC the preliminary raw data ingested. If no QC has been applied, an X will display in the column.

The Last Ingested column lists the date and time data was last received. The Status column identifies the status of data ingested at the sensor.

- Green indicates data from the last ingest is present.
- Blank field indicates there is no data present
- Red X indicates bad data
- Yellow exclamation indicates that data collection has not been attempted

Add a Sensor

Select the Add button to add a sensor. The Add Sensor pop up contains fields for the sensor's metadata on the Sensor Information tab, shown in Figure 3. Required fields are marked with an asterisk (*).

Click the **Timezone** dropdown menu to view and select the sensor's timezone. Customize the **Ingest Frequency** in minutes by entering a numeric value for the frequency.

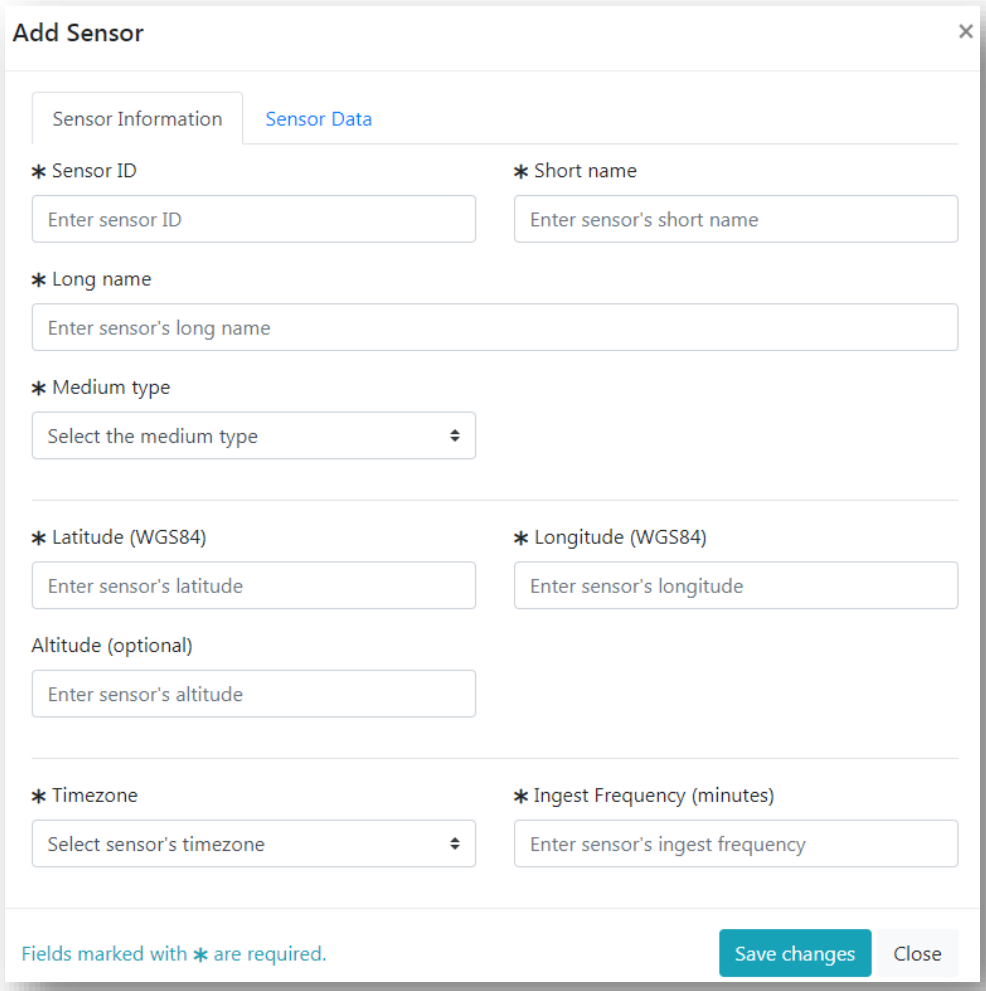

Figure 3. Add Sensor form – Sensor Information tab

The Sensor Data tab (Figure 4) on the Add Sensor pop up contains fields defining the data retrieved from the sensor. Define the data location by entering a URL. Select the Data Quality (raw, preliminary, or final) from the dropdown menu and enter a numeric value to assign the **Timestamp Data Column** . Apply quality control (QC) to the raw or preliminary parameter data by selecting the checkbox. Quality Control matching the selected parameter will be applied to the preliminary raw sensor data ingested. Values that do not meet the QC measure are discarded.

Add multiple parameters to the sensor by selecting a parameter, unit and data column and click Add. Each parameter will be listed in the **Parameter Data Columns** grid (Figure 4).

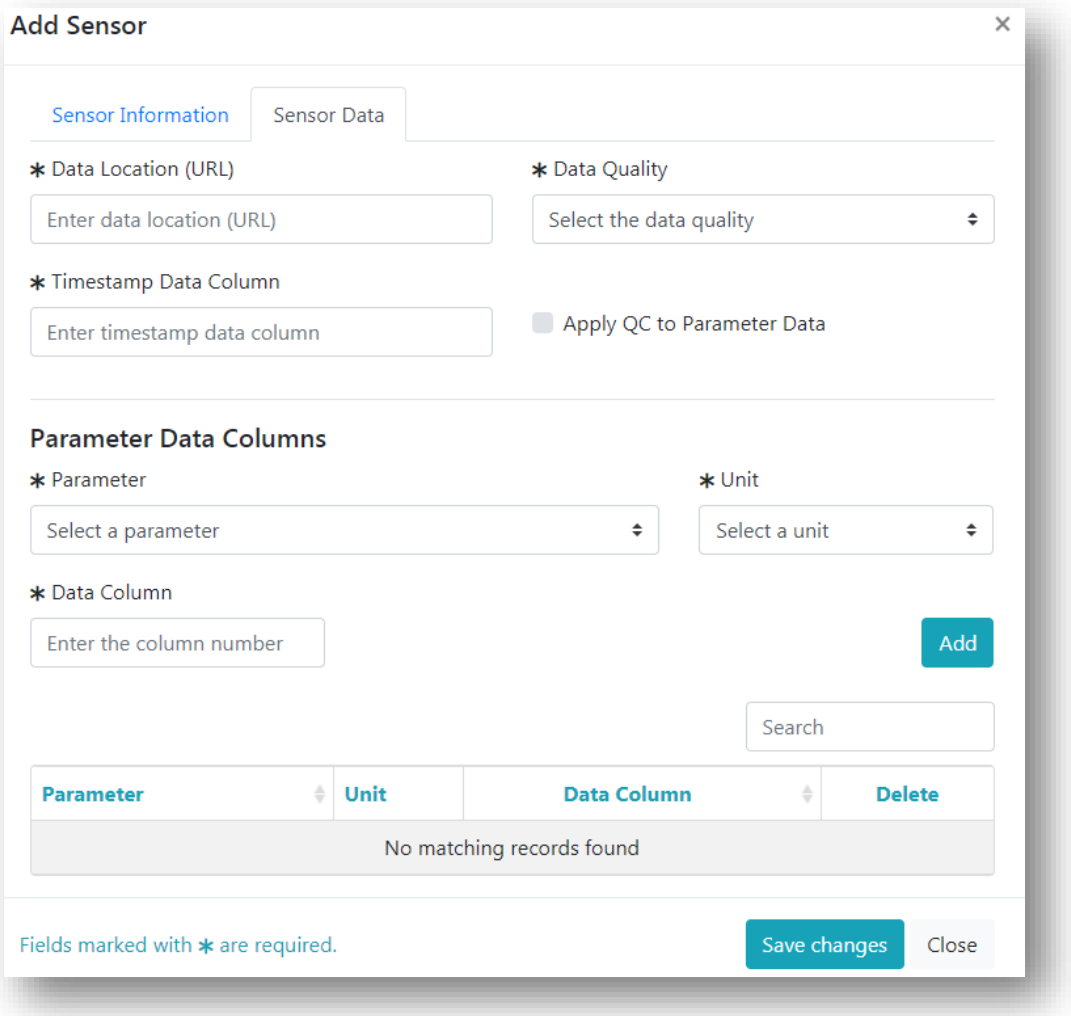

Figure 4.Add Sensor form – Sensor Data tab

Edit a Sensor

Edit a sensor by clicking the edit icon in the Sensors grid. Remove a sensor by clicking the icon.

The **Edit Sensor** pop up displays previously entered information and sensor data for the selected sensor. Changes can be made to all fields on the **Sensor Information** tab except for **Sensor ID** as shown in Figure 5.

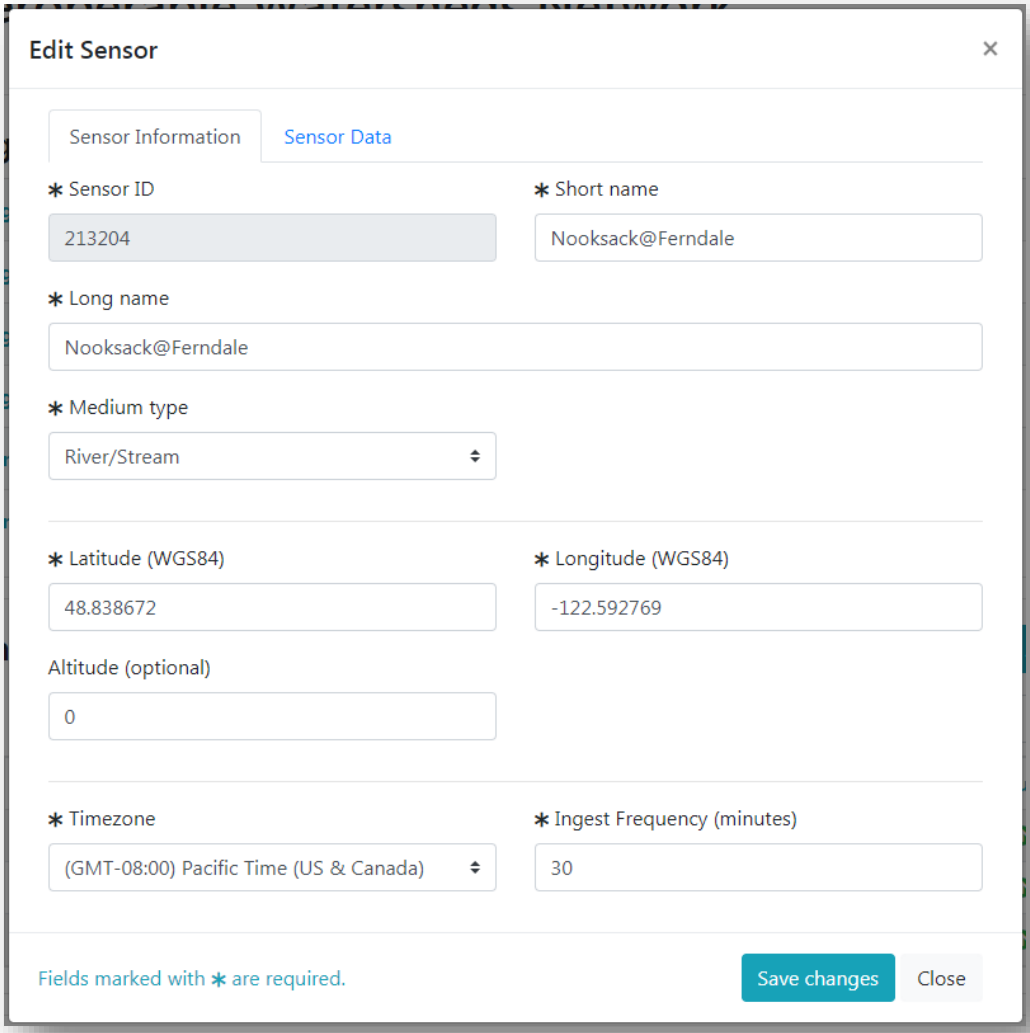

Figure 5. Edit sensor form - Sensor information tab

You can edit a previously entered sensor data on the **Sensor Data** tab of the **Edit Sensor** pop up (Figure 6). Remove a parameter from the sensor by selecting the Delete checkbox for the selected parameter in the **Parameter Data Columns** grid and click **Save changes** shown in Figure 6. Select **Close** or the "**X**" button to exit the **Edit Sensor** pop up without saving any changes made.

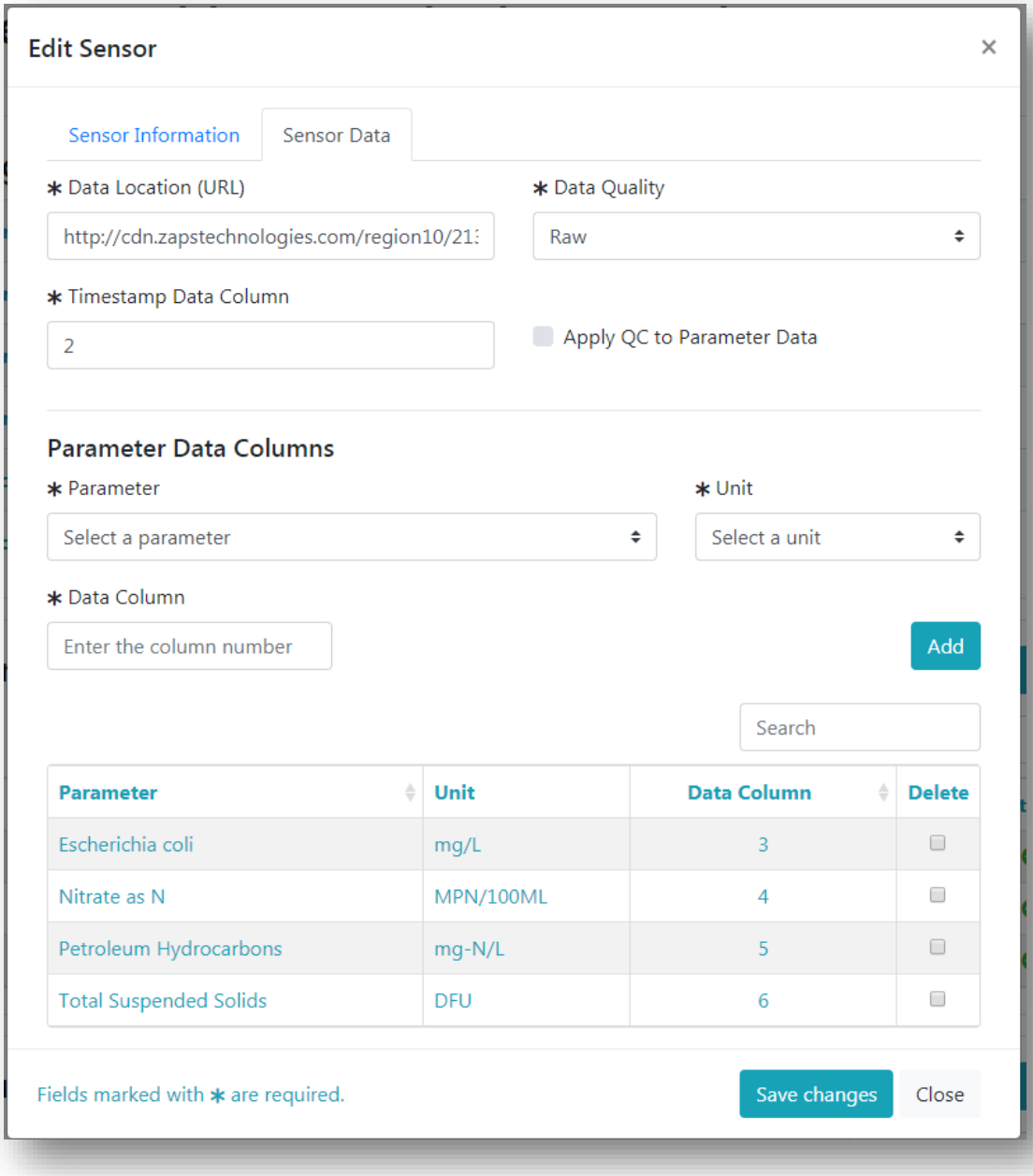

Figure 6. Edit sensor form - Sensor data tab

Quality Control

The Quality Control grid displays each quality control measure defined. Select a column header to sort the grid by parameter, operand, threshold or action. Filter the Quality Control grid by entering text in

the Search field. The grid automatically updates to display filter results. Clear the Search field to reset the Quality Control grid.

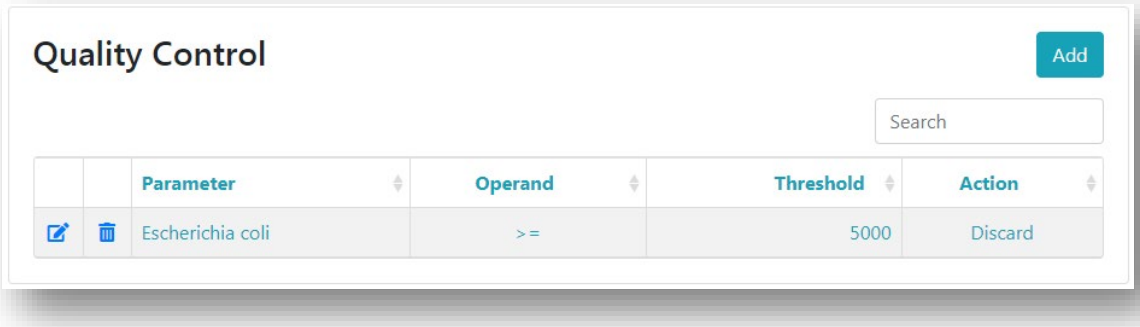

Figure 7. Quality Control grid

Add Quality Control

Select the **Add** button to define and add a new quality control. The **Quality Control** pop up, shown in Figure 8, allows you to select a **Parameter** from the dropdown menu. Set an **Operand** using the dropdown menu. Available operands are listed in the table below.

=	Equal to
<	Less than
<=	Less than or equal to
>	Greater than
>=	Greater than or equal to

The selected operand is applied to the **Threshold**. Enter a numeric value for the **Threshold**. Choose an **Action** to be applied by the quality control.

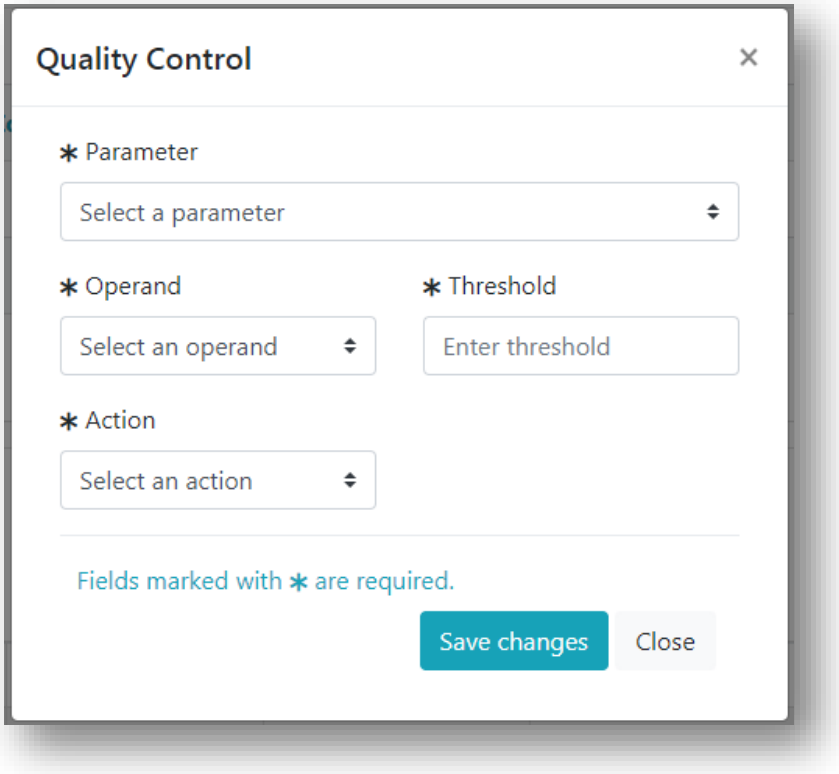

Figure 8. Add Quality Control form

Edit Quality Control

Edit a quality control measure by clicking the edit icon in the **Quality Control** grid. Remove a quality control by clicking the icon.

Changes can be made to all fields in the **Quality Control** pop up shown in Figure 9. Select **Save changes** to apply and save changes made to the quality control. Select **Close** or the "**X**" button to exit the **Quality Control** pop up without saving any changes made.

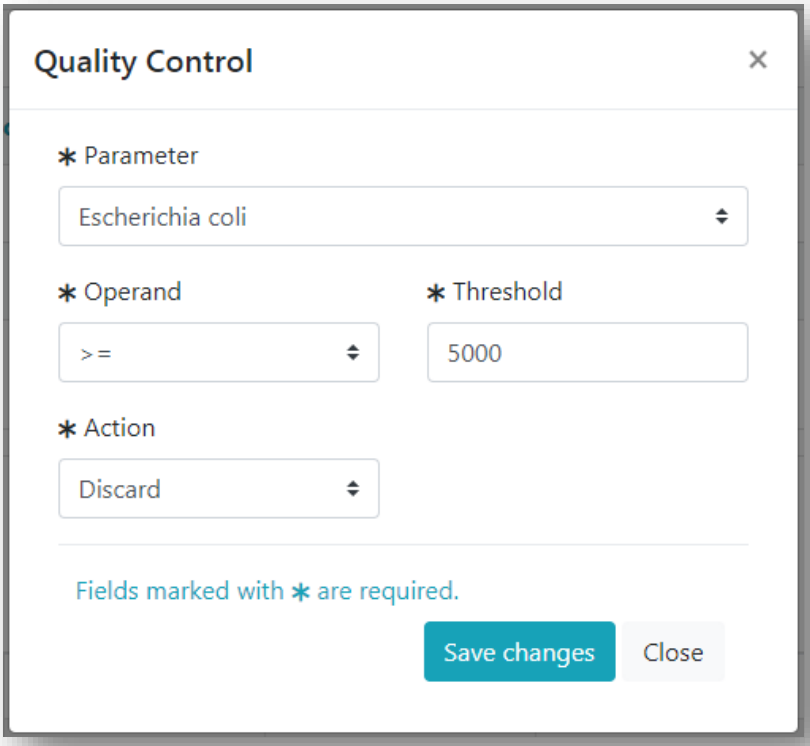

Figure 1. Edit Quality Control form