

Metropolitan Syracuse Wastewater Treatment Facility Data Integrator Notes

This site is a ADG system, operated by OC Department of Water Environment Protection which includes one (1) 350 kW Caterpillar Engine. Thermal output from the unit will be used to heat two of the digesters on the premises. Data collection and monitoring for this site is completed by CDH Energy Corp.

Data Point Details

The data at this site is collected through Onondaga Metro's own SCADA system. Sample for each data channel are taken on the order of once per second, and 1,440 1-minute data records are produced for the raw data collected from this site. This raw data is summed, averaged, or the maximum value is taken for each set of 60-records, resulting in the hourly data uploaded to the Data Integrator website. The resulting hourly data is uploaded on a nightly basis. The details for each individual data point are outlined below.

The timestamp in the raw data files is in Eastern Standard Time. All graphical figures on the website are presented in Eastern Standard Time. This means that during the Daylight Savings Time period from the first Sunday in April until the last Sunday in October the monitored data plots, CSV output and standardized PDF reports are in Eastern Standard Time and do not obey Daylight Savings time rules. Presenting data in Standard Time throughout the year is common practice for graphical time series plotting because it eliminates skipping an hour in April and duplicating an hour in October.

Data Quality Checks

The Data Quality Checks consist of three levels of verification: does the data exist, does the data pass reasonable range checking and does the data pass relational checks. The methodology for applying the data quality begins by creating a contiguous database. This is necessary to maintain compatibility between the many sites on the server. Next, the data received for this site is fit into the database, in this case we are using 1-minute data. For any period where there is data, the data quality level is set to 3 for "Passes Relational Checks". We then work backwards to identify data that does not meet Relational and/or Range Checking.

The next step is to apply the relational checks. Relational checks attempt to identify data which is uncorroborated by the rest of the data set. For instance, data received indicating a DG/CHP Generator output when the gas use is zero is suspect. For data failing a relational check, the data quality level is set to 2 for "Data Passes Range Checks" or 1 for "Data Exists".

The last step is evaluating the range checks. The range checks consist of reasonable high and low values based on facility and DG/CHP Generator information. Data that falls outside the defined range for the database value has its data quality level set to 1 for "Data Exists."

It is necessary to work backwards when applying data quality checks to insure that data gets set to the lowest applicable data quality level. It is possible for data to pass the relational check and fail the range check and such data will be set to a data quality level of 1 for “Data Exists.”

Table 1. Data Quality Definitions

Data Quality Levels	Description	Definition
3	Passes Relational Checking	This data passes Range Checks and Relational Checks. This is the highest quality data in the data set.
2	Passes Range Checks	This data passes the Range Checks but is uncorroborated by Relational Checks with other values.
1	Data Exists	This data does not pass Range Checks. This data is found to be suspect based on the facility and/or CHP equipment sizing.
0	Data Does Not Exist	This data is a placeholder for maintaining a contiguous database only.

Details on the Range and Relational Checks are found below.

Relational Checks

These checks are applied to the 1-minute data before it is converted to hourly data. If any of the 1-minute data points fails the relational check, the data for the entire hour is marked as failed. When there is a failure to obtain new data, the data set repeats the old value. We can identify this bad data through a relational check for repeating data on the equipment separately. We are using a threshold of 95% repeating values because some values can reset to zero during the repeating periods.

Table 2. Relational Checks for Onondaga Metro

Evaluated Point	Criteria	Result
FG	$WG > 10$ and $FG \leq 0$	DQ Level for FG set to 2
WG, WG_KW, FG	> 95% of columns repeat previous data record	DQ Level for WG, WG_KW, and FG set to 1

Notes: FG – DG/CHP Generator Gas Use
 WG – DG/CHP Generator Energy
 WG_KW – DG/CHP Generator Power

Range Checks

These checks are applied to the 1-minute data before it is converted to hourly data. If any of the 1-minute data points fails the range check, the data for the entire hour is marked as failed. When there is a failure to obtain new data, the data set repeats the old value.

Table 3. Range Checks for Onondaga Metro

Data Point	Hourly Data Method	Upper Range Check	Lower Range Check
WG (Generator Output)	Sum	375 kWh	0 kWh
FG (Generator Fuel)	Sum	7,500 cubic feet	0 cubic feet
DG/CHP Biogas Valve	Status / Minute	1	0
DG/CHP Natural Gas Valve	Status / Minute	1	0
TAO (Ambient Temperature)	Average	130°F	-30°F

Notes: Data failing the Range Check has the data quality level set to 1 for "Data Exists"

ASERTTI Protocol Adherence

This site adheres to the ASERTTI Long-Term Monitoring Protocol with the following exception: the Inlet Air Temperature is not measured. For analysis, the outdoor air temperature from a nearby weather station has been substituted. All other required performance parameters are reported in 1-minute averages and sums or can be calculated.

Monitoring Notes

July 21, 2009

Test run performed and first data file received.