**Config.json File Documentation**

**simulation\_metadata\_name**

* File name of the simulation metadata csv file
* Example value: "summary\_results\_algorithm\_Knoxville\_TN\_AMY\_brief\_2.csv"

**weight\_of\_baseline\_data**

* Since the simulation data is imbalanced, meaning that the non-fault data accounts for only a small portion of the whole dataset. This configuration is a multiplier to duplicate the current non-fault data records
* Example value: 1

**random\_state**

* Random seed for all random process in the code
* Example value: 42

**number of trees**

* Number of trees for random forest or gradient boosting algorithms
* Example value: 25

**considered\_sensors\_in\_sensor\_inaccuracy\_analysis**

* Subset list of sensors to be considered in sensor inaccuracy analysis, which is a list of string
* Example value: ["cooling\_electricity [W]", "electricity\_facility [W]", "whole\_building\_facility\_total\_hvac\_electric\_demand\_power [W]"]

**sensor\_category\_dict**

* A dictionary to map each sensor name and their sensor type. This is used in sensor inaccuracy analysis
* Example value: {"cooling\_electricity [W]": "electiricty\_meter",

"electricity\_facility [W]": "electiricty\_meter", "whole\_building\_facility\_total\_hvac\_electric\_demand\_power [W]": "electiricty\_meter"}

**sensor\_fault\_probability\_table\_dir**

* The csv file name of the sensor fault probability table. The details of fault probability table can be found in “documentations/Fault\_Probability\_Talbe\_Description.docx”
* Example value: "sensor\_fault\_probability\_table.csv"

**selected\_fault\_types\_for\_sensor\_inaccuracy\_analysis**

* Subset of fault types to be considered for sensor inaccuracy analysis, which is a list of string.
* Example value: ["air\_handling\_unit\_fan\_motor\_degradation",

"biased\_economizer\_sensor\_mixed\_t",

"duct\_fouling"]

**sensor\_group\_info\_dir**

* The csv file name of the sensor group information. The sensors are grouped into basic, moderate, rich, and unlikely. The details of fault probability table can be found in “documentations/Sensor\_Group\_Info\_Description.docx”
* Example value: "Suggested Sensor Sets.csv"

**fault\_type\_list\_sensor\_cost\_analysis**

* A subset of fault type to be considered in the sensor cost analysis.
* Example value: ["thermostat\_bias", "economizer\_opening\_stuck"]

**baseline\_sensor\_set\_sensor\_cost\_analysis**

* Baseline sensor set in the format of list of string, which is used for the settings in sensor cost analysis. It defines the existing sensors that already exist in the building.
* Example value: [“room\_103\_zone\_mean\_air\_temperature [C]”,

“room\_104\_zone\_mean\_air\_temperature [C]”]

**candidate\_sensor\_set\_sensor\_cost\_analysis**

* Candidate sensor set in the format of list of string, which is used for the settings in sensor cost analysis. It defines the candidate sensors to be evaluated and installed in the building in the future.
* Example value: ["room\_204\_zone\_air\_relative\_humidity [%]", "electricity\_facility [W]"]

**fault\_prevalence\_dict**

* A dictionary with fault type as keys and the prevalence of fault as values. It is used to aggregate sensor opportunity cost by fault types.
* Example value: {"air\_handling\_unit\_fan\_motor\_degradation": 0.255,

"baseline": 0.05,

"biased\_economizer\_sensor\_mixed\_t": 0.19}

**guidance\_bldg\_dict**

* A dictionary with building name as keys. The values are currently unused and set to dummy values.
* Example value: {"small\_commercial\_building": "None"}

**algorithm\_dict**

* A dictionary with supported machine learning algorithm as keys. The values are currently unused and set to dummy values.
* Example value: {"random\_forest": "None", "gradient\_boosting": "None"}