Project 8 - Buoy Data Generator

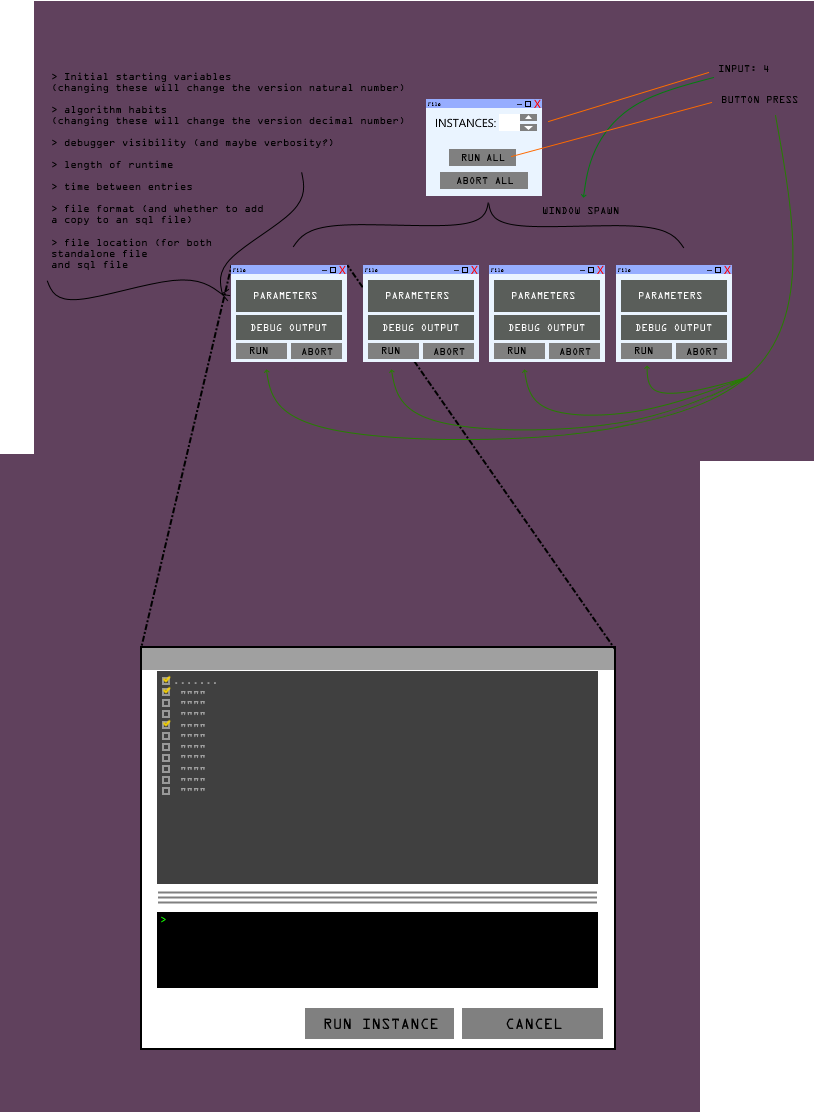
Requirements

The Buoy data generator will generate new, randomised data based on the CSV file specifications of existing NDBC Buoy data files which are in CSV file format. The existing NDBC Buoy data files have previously been converted over from NetCDF file format to CSV file format using a separate program called NetCDF\_to\_csv.py which is currently a Python program run via the command-line.

**User Requirements Story:**

The user has existing NDBC buoy data for e.g. the year 2020 and wants to generate simulated data for the years 2022-2025. The user should:

1. Run/Start the Buoy Data Generator application
2. Click “Select Config File” button
3. Browse for the desired Config file (YAML formatted file), select it and click ‘OK’
4. Use the up/down arrow selectors to choose the number of data generation instances required
5. Click ‘Run All’ to run the data generation process based on the selected configuration file



**Configuration File Requirements:**

|  |  |
| --- | --- |
| **---** | **DESCRIPTION:** |
| Theme: |  |
| Background: 0x555555 |  |
| Forground: 0x888888 |  |
| Highlights: 0xaaaaaa |  |
| text: 0xdddddd |  |
| Dimensions: |  |
| X: null |  |
| Y: null |  |
| Position: |  |
| X: null |  |
| Y: null |  |
| Defaults: |  |
| Dark theme: True |  |
| History: null |  |
| BASE\_FOLDER: C:/Users/afox/Downloads/cmanwx |  |
| BASIS\_YEAR: '2020' |  |
| BASIS\_MONTH: '01' |  |
| CSV\_FOLDER: 'csv' | Input data is in: BASE\_FOLDER/year/month/csv/... |
| MAX\_FILES\_TO\_RUN: 1000 | The maximum number of files/buoys to generate data |
| OUTPUT\_YEARS: |  |
| [2022, 2023, 2024, 2025] | List of the years desired for the output sim data |
| OUTPUT\_MONTHS: |  |
| [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] | List of the months desired for the output sim data |
| MonthLength: |  |
| [0,31,28,31,30,31,30,31,31,30,31,30,31] |  |
| Statscolumns: |  |
| - FieldName | Name of the field |
| - DataType | Type of the field (int, float, string, datetime) |
| - Min | Numerical minimum |
| - Max | Numerical maximum |
| - Mean | Numerical average (mean) |
| - StdDev | Standard deviation |
| - Median | Field median value |
| - Mode | Most likely/common value of this field |
| - NumValues | Total number of values (including nulls) |
| - NumNulls | Number of null values |
| - NumUnique | Number of unique values for this field |
| - AutoCorr | Field self-correlation |
| - FFT | Data periodicity (hrs) |
| - Distrib | Distribution (discrete only) |

frequency of the data

any mandatory fields