CLAIMS

1. A method, comprising:

providing first data comprising a random or pseudo-random input to a first trained machine learning model trained to generate synthetic sensor time series data for a processing chamber;

providing second data indicative of one or more attributes of target synthetic sensor time series data to the first trained machine learning model, the second data comprising one or more labels identifying one or more of a data source of interest comprising a sensor type, sensor location, or processing recipe associated with a processing chamber, or a state of interest of the processing chamber;

receiving an output from the first trained machine learning model based on the first data and the second data, wherein the output comprises synthetic sensor time series data associated with the processing chamber and the data source of interest or the state of interest, wherein the output is generated in view of the second data indicative of one or more attributes; and

utilizing the synthetic sensor time series data to train a second machine learning model to control operation of the processing chamber.

2. The method of claim 1, further comprising training the first machine learning model, wherein training the first machine learning model comprises:

causing the first machine learning model to generate synthetic sensor time series data;

providing the synthetic sensor time series data to a third machine learning model comprising a discriminator;

providing measured sensor time series data to the third machine learning model, wherein the third machine learning model is configured to distinguish between synthetic sensor time series data and measured sensor time series data;

providing feedback data to the first machine learning model, indicative of how accurately the third machine learning model distinguished synthetic from measured sensor time series data; and

updating the first machine learning model to generate synthetic sensor time series data that the third machine learning model less accurately distinguishes from measured sensor time series data.

3. The method of claim 1, wherein the first trained machine learning model comprises a generator of a generative adversarial network.

4. The method of claim 1, wherein the first trained machine learning model comprises a recurrent neural network model.

5. The method of claim 1, wherein the synthetic sensor time series data comprises data corresponding to one or more of:

power, voltage, or current supplied to a component of the processing chamber;

pressure; or

temperature.

6. The method of claim 1, wherein training the second machine learning model comprises:

providing the output synthetic sensor time series data to the second machine learning model as training input; and

providing third data indicative of one or more attributes associated with the output synthetic sensor time series data to the second machine learning model as target output, wherein the second machine learning model is configured to predict attributes of the processing chamber based on measured sensor time series data of the processing chamber.

7. The method of claim 6, wherein the second machine learning model is configured to detect one or more anomalies associated with measured sensor time series data of the processing chamber.

8. The method of claim 1, wherein an attribute of target synthetic sensor time series data comprises one or more of:

time since installation of the processing chamber;

time since a previous maintenance event of the processing chamber; or

a fault present in the processing chamber.

9. A system, comprising memory and a processing device coupled to the memory, wherein the processing device is configured to:

provide first data comprising a random or pseudo-random input to a first trained machine learning model, wherein the first trained machine learning model is trained to generate synthetic sensor time series data for a processing chamber;

provide second data indicative of one or more attributes of target synthetic sensor time series data to the first trained machine learning model, the second data comprising one or more labels identifying one or more of a data source of interest comprising a sensor type, sensor location, or processing recipe associated with a processing chamber or a state of interest of the processing chamber;

receive an output from the first trained machine learning model based on the first data and the second data, wherein the output comprises synthetic sensor time series data associated with the processing chamber and the data source of interest or the state of interest, wherein the output is generated in view of the second data indicative of one or more attributes; and

utilize the synthetic sensor time series data to train a second machine learning model to control operation of the processing chamber.

10. The system of claim 9, wherein the processing device is further configured to train the first machine learning model, wherein training the model comprises:

causing the first machine learning model to generate synthetic sensor time series data;

providing the synthetic sensor time series data to a third machine learning model;

providing measured sensor time series data to the third machine learning model,

wherein the third machine learning model is configured to distinguish between synthetic sensor time series data and measured sensor time series data;

providing feedback data to the first machine learning model indicative of how accurately the third machine learning model distinguished synthetic from measured sensor time series data; and

updating the first machine learning model to generate synthetic sensor time series data that the third machine learning model less accurately distinguishes from measured sensor time series data.

11. The system of claim 9, wherein the first trained machine learning model comprises a generator of a generative adversarial network.

12. The system of claim 9, wherein the first trained machine learning model comprises a recurrent neural network model.

13. The system of claim 9, wherein the synthetic sensor time series data comprises data corresponding to one or more of power, voltage, or current supplied to one or more of:

a radio frequency plasma generation component;

a heater; or

a substrate support,

pressure; or

temperature.

14. The system of claim 9, wherein the processing device is further configured to:

train a second machine learning model, wherein training the second machine learning model comprises:

providing the output synthetic sensor time series data to the second machine learning model as training input; and

providing third data indicative of one or more attributes associated with the output synthetic sensor time series data to the second machine learning model as target output, wherein the second machine learning model is configured to predict attributes of the processing chamber based on measured sensor time series data of the processing chamber.

15. The system of claim 14, wherein the second machine learning model is configured to detect one or more anomalies associated with measured sensor time series data of the processing chamber.

16. The system of claim 9, wherein an attribute of target synthetic sensor time series data comprises one or more of:

time since installation of the processing chamber;

time since a previous maintenance event of the processing chamber; or

a fault present in the processing chamber.

17. A non-transitory machine-readable storage medium storing instructions which, when executed, cause a processing device to perform operations comprising:

providing first data comprising a random or pseudo-random input to a first trained machine learning model trained to generate synthetic sensor time series data for a processing chamber;

providing second data indicative of one or more attributes of target synthetic sensor time series data to the first trained machine learning model, the second data comprising one or more labels identifying one or more of a data source of interest comprising a sensor type, sensor location, or processing recipe associated with a processing chamber or a state of interest of the processing chamber;

receiving an output from the first trained machine learning model based on the first data and the second data, wherein the output comprises synthetic sensor time series data associated with the processing chamber and the data source of interest or the state of interest, generated in view of the second data indicative of one or more attributes; and

utilizing the synthetic sensor time series data to train a second machine learning model to control operation of the processing chamber.

18. The non-transitory machine-readable storage medium of claim 17, the operations further comprising:

training the first machine learning model, wherein training the model comprises:

causing the first machine learning model to generate synthetic sensor time series data;

providing the synthetic sensor time series data to a third machine learning model;

providing measured sensor time series data to the third machine learning model, wherein the third machine learning model is configured to distinguish between synthetic sensor time series data and measured sensor time series data;

providing feedback data to the first machine learning model, indicative of how

accurately the third machine learning model distinguished synthetic from measured sensor time series data; and

updating the first machine learning model to generate synthetic sensor time series data that the third machine learning model less accurately distinguishes from measured sensor time series data.

19. The non-transitory machine-readable storage medium of claim 17, wherein an attribute of target synthetic sensor time series data comprises one or more of:

time elapsed since installation of the processing chamber;

time elapsed since a previous maintenance event of the processing chamber; or

a fault present in the processing chamber.

20. The non-transitory machine-readable storage medium of claim 17, the operations further comprising:

training a second machine learning model, wherein training the second machine learning model comprises:

providing the output synthetic sensor time series data to the second machine learning model as training input; and

providing third data indicative of one or more attributes associated with the output synthetic sensor time series data to the second machine learning model as target output, wherein the second machine learning model is configured to predict attributes of the processing chamber based on measured sensor time series data of the processing chamber.