Predicting Atmospheric Greenhouse Gas Concentration Using Gaussian Processes

Suleiman Adamu Yakubu (11796778)

Eziamaka Nwakile (11990516)

Coventry University

Modeling and Optimization Under Uncertainty (7135CEM)

Dr. Alireza Daneshkhah

26/August/2022

Abstract

This paper intends to investigate the atmospheric concentration of greenhouse gasses (CO_2, N2O, CH_4, and SF_6) and how the same has changed over the years. This phenomenon is directly linked to global warming. The paper then attempts to predict the concentration of these gases over the next 36 months using Gaussian Processes (GPs) on a dataset obtained from the Mauna Loa Observatory. The paper also compares the performance of GPs with some other regression models. These regression models were picked as the best candidates for the dataset using Pycaret. The Gaussian Process models outperformed the other models and returned a confidence interval on their predictions for each gas.

Keywords: Gaussian Processes, machine learning, regression, sklearn, pycaret, global warming, greenhouse gas.

Introduction

The earth is becoming warmer. This has been observed since the pre-industrial era (1850-1900) (National Aeronautics and Space Administration, 2020). While the temperature of the earth has never stayed constant, higher rates of change are now being observed and it is generally agreed that this rise is due to human activity, mainly, the production of the so-called greenhouse gasses (Hegerl et al., 1996, 2291-2306). These gasses contribute to the increase in the earth's temperature through the greenhouse effect (hence their name). They are found in the atmosphere after they are released through human activity (mainly the burning of fossil fuels). They envelop the earth, and let heat from the sun through to hit the earth, however, they do not allow as much reflected heat from the earth to escape. This leads to the surface temperature of the earth rising.

As man's activities, especially energy generation, lead to higher and higher concentrations of the greenhouse gasses in the atmosphere (about 25%), what is needed according to most observers is to limit the temperature rise to at most 2ºC above pre-industrial levels. This can only be done if the atmospheric concentration of greenhouse gasses (especially CO_2) is maintained between 350ppm and 450ppm (Bakay & Agbulut, 2021).

This gives rise to the question, at the rate at which the concentration of the gasses are increasing now, how much time will it take before it reaches the threshold? There are various methods to predict this.

(Bakay & Agbulut, 2021) used Support Vector Machines and Deep Learning to forecast greenhouse gas emissions. They however limited themselves to greenhouse gas emissions due to electrical energy generation in Turkey. Furthermore, while the models used were considered highly accurate, the models do not provide a confidence level for their forecast

(Javanmard & Ghaderi, 2022) proposed and implemented a hybrid model that used machine learning algorithms and an optimization model to forecast greenhouse gas emissions (CO_2, N2O, CH_4, and Fluorinated-Gas). They used energy data to do so. They managed to gain increased accuracy while using the optimization models. They however restricted their forecasting to the greenhouse gas emissions of Iran. Their models also did not provide a confidence level for the forecasts.

Finally, (Ho & Yu, 2022) forecasted the greenhouse gas emissions of Hong Kong using log-linear regression models. They also tried to select optimal predictors for the forecast. As such, their model included principal components such as social and economic factors. Their model achieved a CV(RMSE) of 6.1508% in the testing sets. Again, they focused on the greenhouse emissions of one country (Hong Kong) and did not provide confidence levels for their predictions

Gaussian Processes, which this report focuses on, can be used to forecast as well. They have the advantage of being non-parametric, and provide a confidence level for each prediction they make. As such, they can be used where data is limited. The confidence level for the prediction is also important because decisions taken using their predictions can then factor how sure they are that the predictions are accurate.

Problem Statement and Dataset

The problem that this report and project aims to solve is the problem of information. The project aims to project into the future and predict what the concentrations of the greenhouse gasses are if their production continues on the current trajectory. It can also help us to see how much time is available before the concentration reaches the threshold. This helps policy makers make appropriate plans in the battle for global warming.

The dataset used is obtained from the Mauna Loa Observatory in Hawaii. The dataset contains 4 files (one for each gas). Each file contains the following:

Variable Name	**Variable Format**	**Description**
Year	Integer	The year the measurement was taken
Month	Integer	The month the measurement was taken
Decimal Date	Decimal	The year and month converted to decimal format
Monthly Average	Decimal	The average concentration measured for that month
Deseasonalized average	Decimal	The average measurement deseasonalized
Days	Integer	
Standard deviation of days	Decimal	Standard deviation of days
Uncertainty	Decimal	The uncertainty of the measurement

Methodology

Gaussian Process Regression

Supervised learning in machine learning is concerned with learning patterns in data where the correct output/label is attached to the data. This means that during training, the model knows what the label is meant to be. The data is given in the form $D = \{(x_1, y_1), ..., (x_n, y_n)\}$ with x being a vector of observed points and y being the vector of labels. Gaussian Processes can be used to make predictions in a supervised learning context.

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution (Williams & Rasmussen, 2006). Gaussian Processes are useful for prediction because they define a distribution over functions. This means that given a Gaussian Process (GP), which is specified by its mean and covariance function, at any point x we can sample a function according to the following relation

$$f(x) \sim GP(\mu(x),\ k(x,\ x'))$$

Where $f(.)$is the function sampled from the GP, $\mu(.)$ is the mean function and $k(.,.')$ is the covariance function (usually represented as Σ). Note that the covariance function specifies the similarities between the points.

For Gaussian Process Regression (GPR), we set the mean function to zero and covariance function $k(x,\ x')$ (also called the kernel) to the following:

$$k(x,\ x') = E[(f(x) - \mu(x))(f(x') - \mu(x'))] \in \mathcal{R}$$

Although it is said that GPs are non-parametric, they are actually parametrized by the mean function $\mu(x)$ and the kernel $k(x,\ x')$. The kernel ensures that points that are similar (i.e. close together in the input space) will produce outputs that are similar. How the GPR makes predictions is then as follows: given a dataset $D = \{(x_1, y_1),\ ....,\ (x_n, y_n)\}$, and test points x_*. We specify the prior to be a Gaussian ($Prior() =$. The GPR then returns a distribution of functions that fits the dataset.

Covariance Matrices (Kernels)

There are different possibilities to use when specifying the kernel for a Gaussian Process Regressor (GPR). The general form of a kernel or covariance matrix Σ is given as:

$$\begin{bmatrix} K & K_* \\ K^T_* & K_{**} \end{bmatrix}$$

The different kernels used with GPs are as follows:

1. **Radial Basis Function (RBF) Kernel**: The RBF kernel helps to model smoothness in the data and is expressed as: $k(x, x') = \sigma^2 exp\left(-\frac{(x-x')^2}{2l^2}\right)$. The main parameters of this kernel are the variance σ^2 and the length scale l. Optimizing these will result in better predictions.
2. **Exponential Sine Squared Kernel**: This kernel helps to model seasonality or periodicity in the data. It is expressed as: $k(x, x') = \sigma^2 exp\left(-\frac{2}{l^2}sin^2\left(\pi\frac{|x-x'|}{p}\right)\right)$. The parameters are the variance σ^2, the length scale l, and the period p which is the distance between repetitions.
3. **Rational Quadratic Kernel:** This kernel has a similar function to the RBF kernel. They both model smoothness. The kernel is given as: $k(x, x') = \sigma^2(1 - \frac{||x-x'||^2}{2\alpha l^2})^2$. The parameters of this kernel are σ^2, α, and l which are the variance, scale mixture ($\alpha >$ 0) , and length scale.
4. **White Noise Kernel:** This kernel is used to model noise in the data. It can be expressed as: $k(x, x') = \sigma^2 I_n$ Where I_n is an identity matrix with dimensions $n \times n$. This is because the noise is uncorrelated.

Because all these kernels model Gaussian priors, multiplying, adding, marginalizing, and conditioning them will always produce another Gaussian. This property allows for the kernels to be combined as needed to accurately model the data.

Experimental Setup

The code and experiments performed were implemented in Google Colab. This is a free environment that allows Python to be run in cells. The environment also comes with some essential libraries such as NumPy and Pandas preinstalled. The libraries used for this project include the libraries mentioned above and sklearn for machine learning, plotly, and matplotlib for visualizations.

The dataset did not require preprocessing and has no missing values. Although no data preprocessing was done, the data was explored and visualized to get a good understanding of it. This enabled good choices to be made when picking the kernels used for the actual regression.

Visualizing the Data

The following plots show what the data looks like and what the overall trend is.

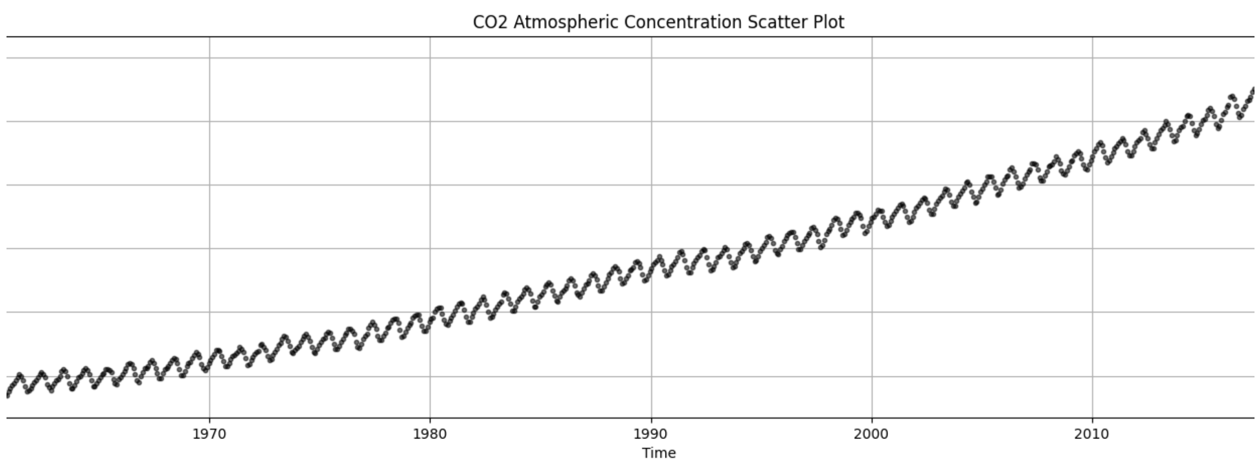


Figure 1: Scatter plot for CO2 monthly concentrations

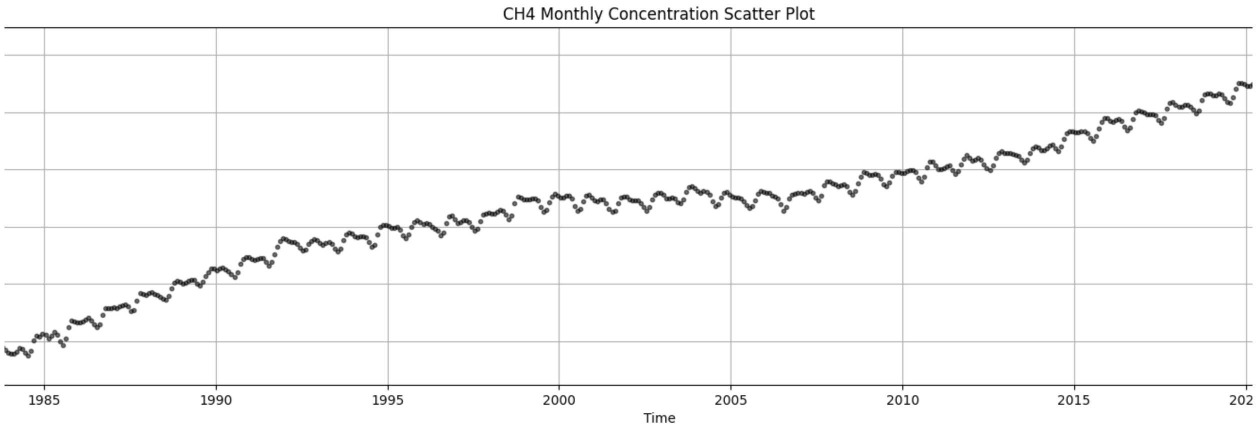


Figure 2: The scatter plot for CH4

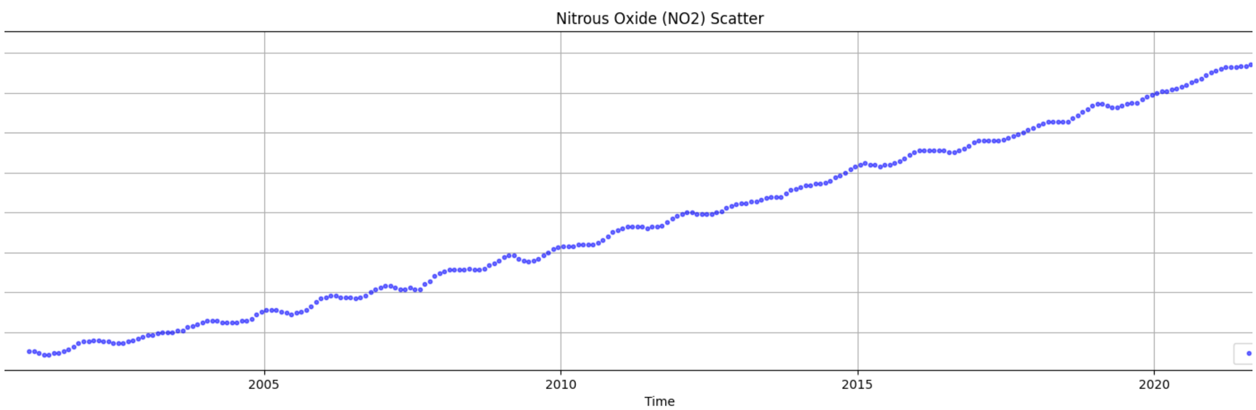


Figure 3: The scatter plot for N2O

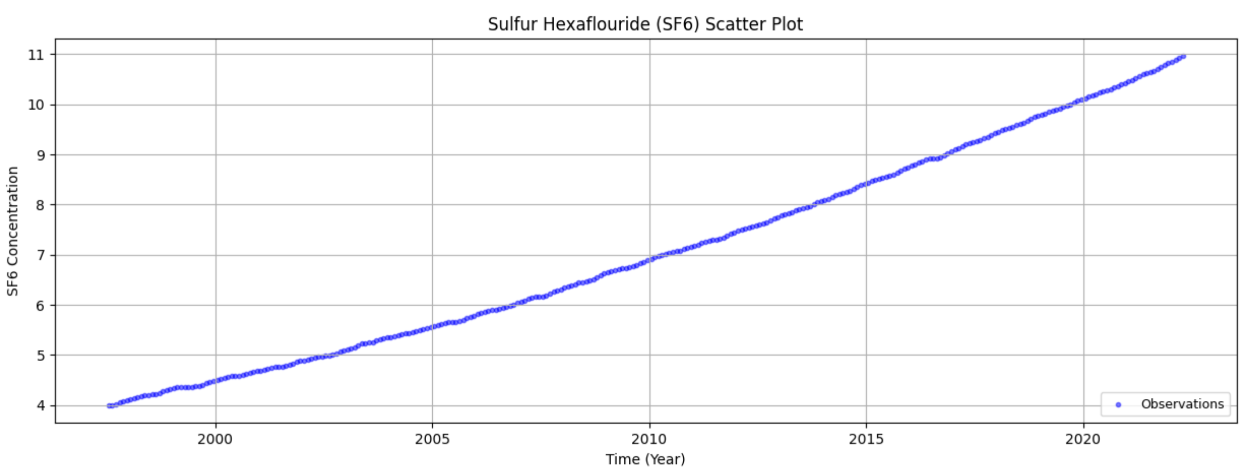


Figure 4: The scatter plot for SF6

For each gas visualized above, it is quite clear that as time passes, the concentration increases. The trend follows a smooth curve and, for some of the gasses, displays some periodicity. This periodicity is subject to slight perturbations, meaning they are not perfectly regular. This suggests that we could use an RBF kernel, along with a White Noise and Exponential Sine kernel to accurately model the data. A snippet of the code is shown below:

```
k1_sf6 = 2.0**2 * RBF(length_scale=100.0) *\
    ExpSineSquared(length_scale=1.0, periodicity=1.0,
periodicity_bounds='fixed')
k2_sf6 = 0.1**2 * RBF(length_scale=0.1) +\
    WhiteKernel(noise_level=0.1**2, noise_level_bounds=(1e-5, 1e5))
k3_sf6 = 2.0**2 * RBF(length_scale=100.0) *\
    ExpSineSquared(length_scale=1.0, periodicity=1.0,
periodicity_bounds='fixed')

sf6_kernel = k1_sf6 + k2_sf6 + k3_sf6
```

This code shows the kernel choice for the Sulfur Hexafluoride predictions. This kernel is trained and optimized automatically during the fitting/training process. This means that the hyperparameters are tuned and reset during training.

Training the Models and Making Predictions

After the kernels have been selected and different hyperparameters tried, it was clear the below configuration worked best for the models. The GPR is then trained on the data. Once the GPR is trained, new points are generated and then the monthly average of the gas concentrations is then predicted. The plot for the CO2 prediction is shown in the figure below, the other plots are attached in Appendix B:

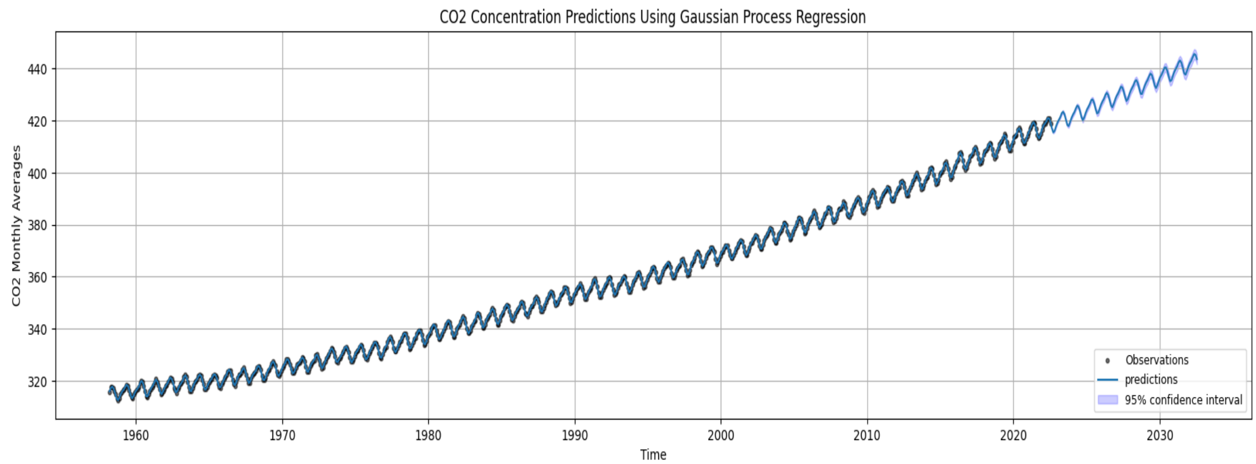


Figure 5: The plot of CO_2 observations and predictions with GPR.

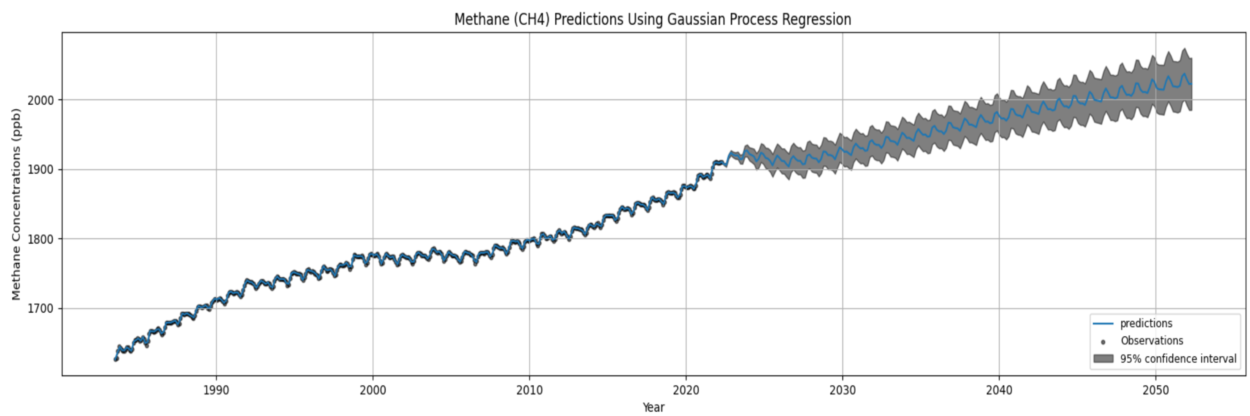


Figure 6: The plot of CH_4 observations and predictions with GPR.

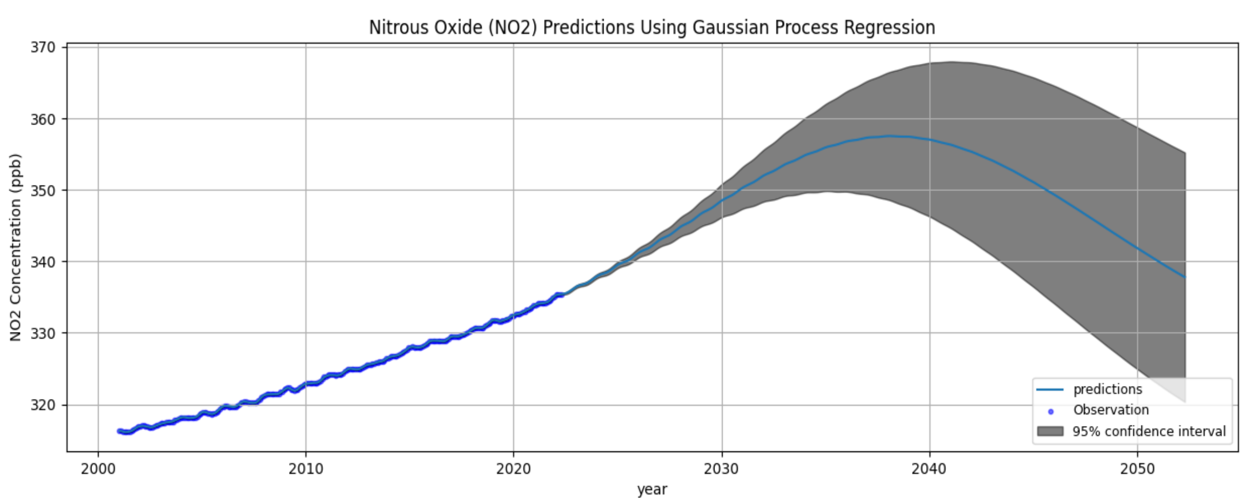


Figure 7: The plot of N2O observations and predictions with GPR.

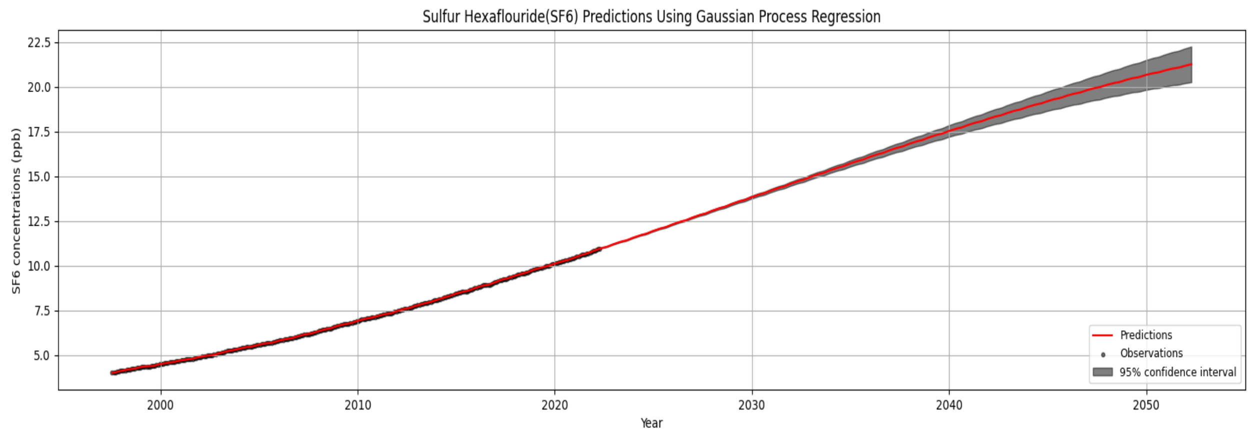


Figure 8: The plot of SF_6 observations and predictions with GPR.

The performance of the GPR was compared against other machine learning models. These models were accessed using the Pycaret library. This library enables a grid search to be done easily to be able to pick the best models for the data and problem at hand. Thus, each gas had different machine learning models used for prediction. The code to create, run, train, and evaluate the models can be found in Appendix A.

One thing all the models had in common was that they performed poorly when predicting on new data. The below figures are some plots of the model predictions. For the rest of the plots, please refer to Appendix B.

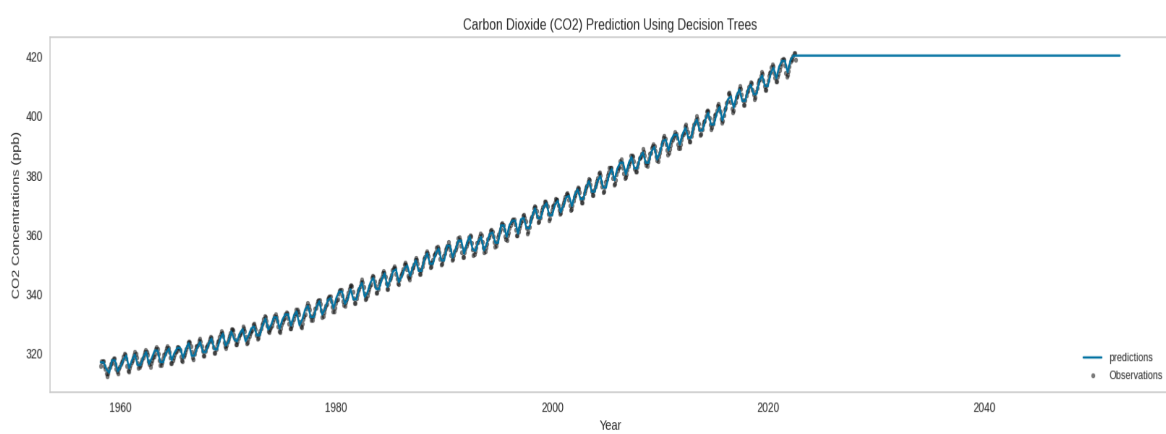


Figure 9: CO2 Concentration prediction using Decision Trees.

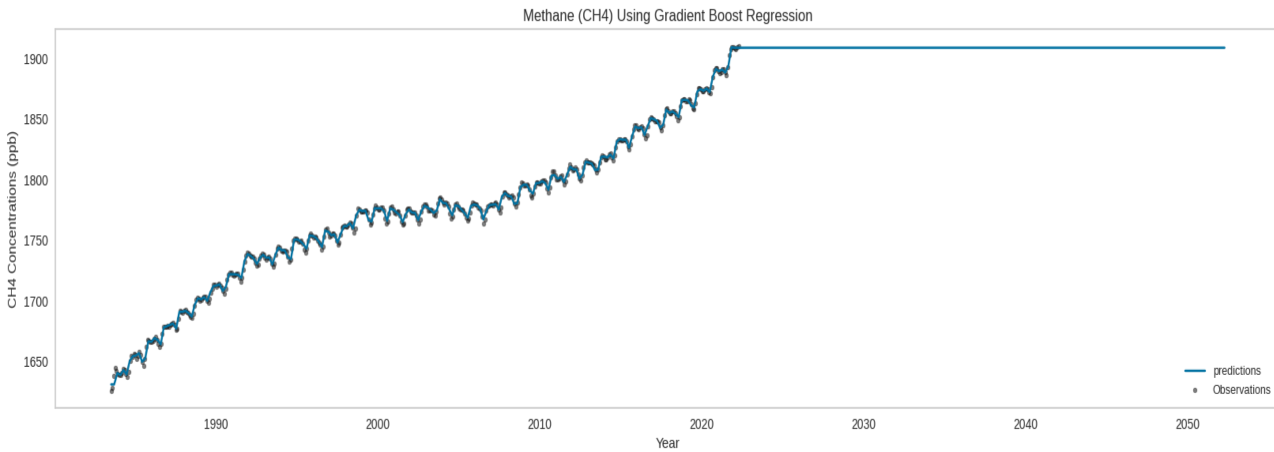


Figure 10: CH4 Concentration prediction using Gradient Boost Regression

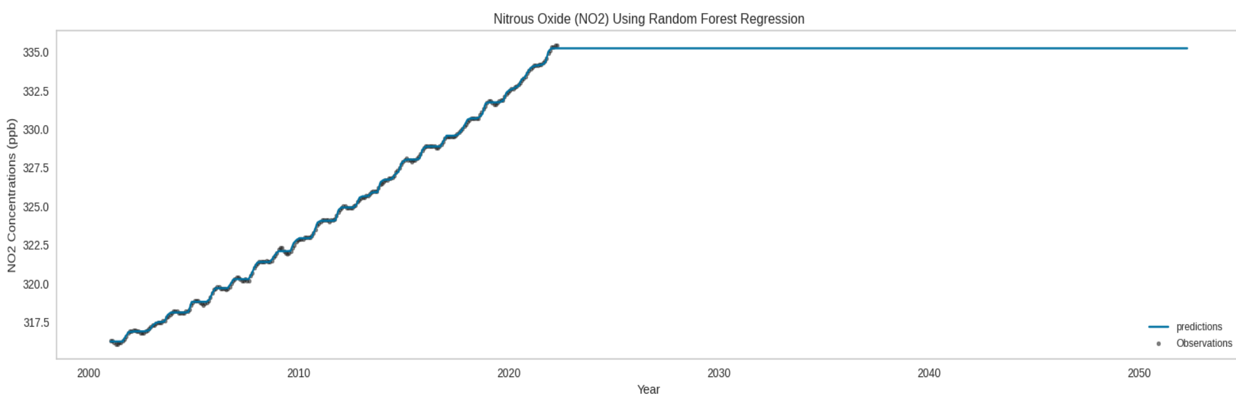


Figure 11: N2O concentration prediction using Random Forest Regression

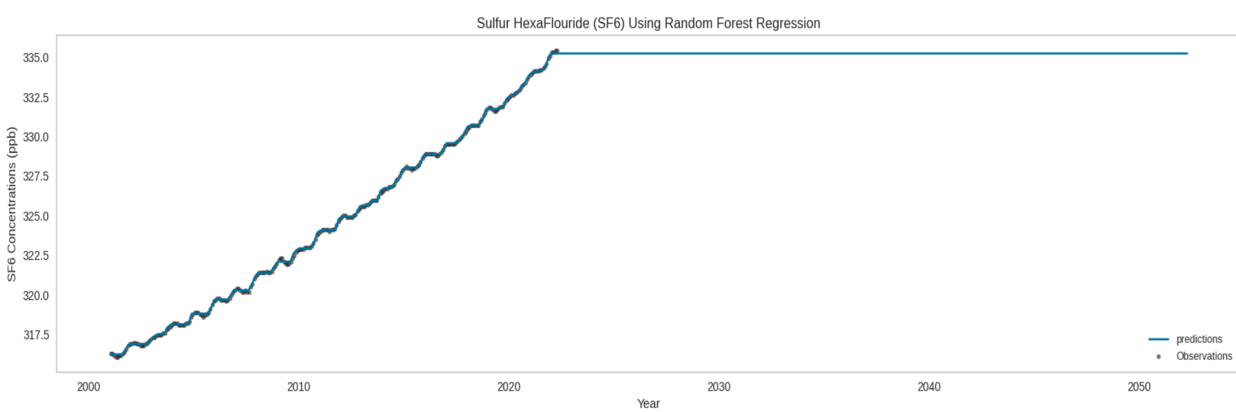


Figure 12: SF6 concentration prediction using Random Forest Regression

Gas	**R^2**	**MSE**	**RMSE**
CO2	0.9999700901489151	0.02748657944621783	0.16579077008753482
CH4	0.9999991557578847	0.003290862124751319	0.05736603633467558
N2O	0.9999785447209445	0.0006733608902757711	0.02594919825882432
SF6	0.9999969149690383	1.2655864061325595e-05	0.0035575081252648736

Table 1: The evaluation results of the GPR

CO2			
Model	***R2***	***MSE***	***RMSE***
Extra Trees	0.9962	3.2675	1.7114
Random Forest	1.8957	1.5398	3.7986
Decision Tree	0.9945	4.7321	2.0994
CH4			
Extra Trees	0.9936	20.6403	4.4014
Random Forest	0.9924	24.9564	4.8751
Gradient Boosting	0.9911	29.1713	5.2574
N2O			
Extra Trees	0.9992	0.0233	0.1408
Random Forest	0.9967	0.0994	0.2607
Gradient Boosting	0.9964	0.1080	0.2864
SF6			
Extra Trees	0.9999	0.0004	0.0171
Random Forest	0.9993	0.0025	0.0435
Decision Tree	0.9989	0.0041	0.0565

Table 2: The results from the other machine learning models

Discussion

In this report, the objective is to predict what the concentrations of some greenhouse gasses will be 30 years in the future. While this can be performed using some deep learning methods such as Recurrent Neural Networks (RNN), they typically do not produce a quantifiable confidence level for their predictions. For this reason, Gaussian Process Regression (GPR) was used.

To accurately follow the scientific process, other machine learning regression models were used as a “control” in this experiment. These machine learning models include Random Forest Regressor, Extra Trees Regressor, Decision Trees Regressor, and Gradient Boosting Regressor. These models produced good performance when presented with data within the timeframe of their training (i.e data they had seen previously). As shown by the metrics and plots provided in the results section, they fit the data very closely. However, they struggled when they were presented with new data points and the prediction was incredibly poor.

The GPR however, produced exemplary results as evidenced by the metrics and plots provided above. In addition, the GPR’s provided a confidence interval for their predictions. This means that when factoring each prediction into the planning process, we can state how sure we are of that prediction.

These predictions can be extremely beneficial to planners as they pinpoint how much time there is left before the concentration of each gas reaches critical level at the current rate of production. Furthermore, these models can be relied on to accurately forecast the concentrations of the gasses if the rates of production change. In other words, they can help us model the consequences of decisions made in cutting down the production of greenhouse gasses.

The data used in this report is open for everyone to use. The Mauna Loa Observatory however requires reciprocation and attribution from anyone that uses the data provided by them. In addition, there are no ethical considerations to using the data.

For further work, the gas concentrations can be modeled as a collective. This will allow planners to be able to predict how the reduction or increase in the production of one or more gasses can affect the others.

References

Bakay, M. S., & Agbulut, U. (2021, February 20). Electricity production based forecasting of greenhouse gas emissions in Turkey with deep learning, support vector machine and artificial neural network algorithms. *Journal of Cleaner Production*, *285*. ScienceDirect. https://doi.org/10.1016/j.jclepro.2020.125324

Dlugokencky, E. (2022). *Mauna Loa N2O Monthly Mean Data* [txt file]. Global Monitoring Laboratories. https://gml.noaa.gov/webdata/ccgg/trends/n2o/n2o_mm_gl.txt

Dlugokencky, E. (2022). *Mauna Loa SF6 Monthly Mean Data* [txt file]. Global Monitoring Laboratories. https://gml.noaa.gov/webdata/ccgg/trends/sf6/sf6_mm_gl.txt

Dlugokencky, E. (2022, August 5). *Mauna Loa CH4 Monthly Mean Data* [txt file]. Global Monitoring Laboratories. https://gml.noaa.gov/webdata/ccgg/trends/ch4/ch4_mm_gl.txt

Hegerl, G. C., Von Storch, H., Hasselman, K., Santer, B. D., Cubasch, U., & Jones, P. D. (1996, October 01). Detecting Greenhouse-Gas-Induced Climate Change with an Optimal Fingerprint Method. *Journal of Climate*, *9*(10), 2281-2306. https://doi.org/10.1175/1520-0442(1996)009<2281:DGGICC>2.0.CO;2

Ho, W. T., & Yu, F. W. (2022, October 10). Optimal selection of predictors for greenhouse gas emissions forecast in Hong Kong. *Journal of Cleaner Production*, *370*. ScienceDirect. https://doi.org/10.1016/j.jclepro.2022.133310

Javanmard, M. E., & Ghaderi, S. F. (2022, July). A Hybrid Model with Applying Machine Learning Algorithms and Optimization Model to Forecast Greenhouse Gas Emissions with Energy Market Data. *Sustainable Cities and Society*, *82*. ScienceDirect. https://doi.org/10.1016/j.scs.2022.103886

National Aeronautics and Space Administration. (2020). *Global Warming vs. Climate Change | Facts – Climate Change: Vital Signs of the Planet*. NASA Climate Change. Retrieved August 15, 2022, from https://climate.nasa.gov/global-warming-vs-climate-change/

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perror, M., & Duchesnay, E. (2011, October). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825-2830. https://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html

Tans, P., & Keeling, R. (2022, July). *Mauna Loa CO2 Monthly Mean Data* [text file]. Global Monitoring Laboratories. https://gml.noaa.gov/webdata/ccgg/trends/co2/co2_mm_mlo.txt

Weinberger, K. (2018, July 11). *Machine Learning Lecture 26 "Gaussian Processes" -Cornell CS4780 SP17* [Video]. YouTube. https://www.youtube.com/watch?v=R-NUdqxKjos&t=570s

Williams, C. K. I., & Rasmussen, C. E. (2006). *Gaussian Processes for Machine Learning*. MIT Press.

Appendices

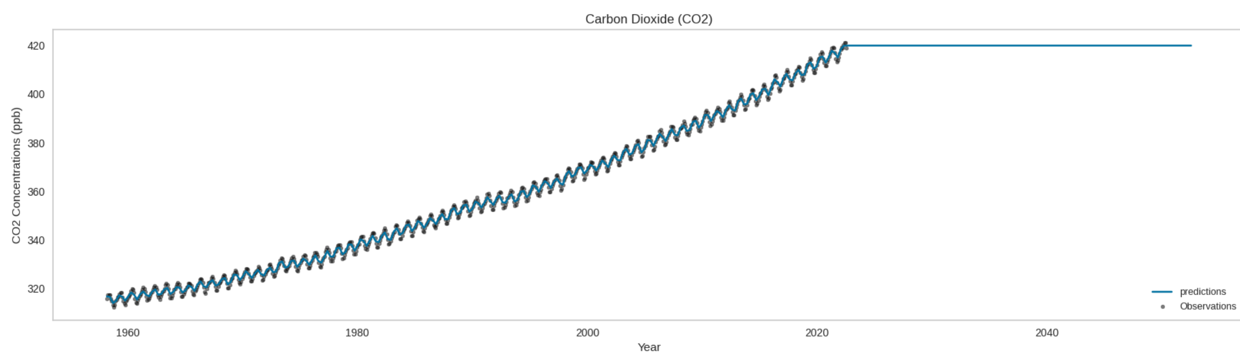
Appendix A (Notebooks with code)

https://colab.research.google.com/drive/1GTUSvo3-E9TZ6B-9Ij9S4Z6cGlWxVPH-?usp=sharing (CO2 GPR Model)

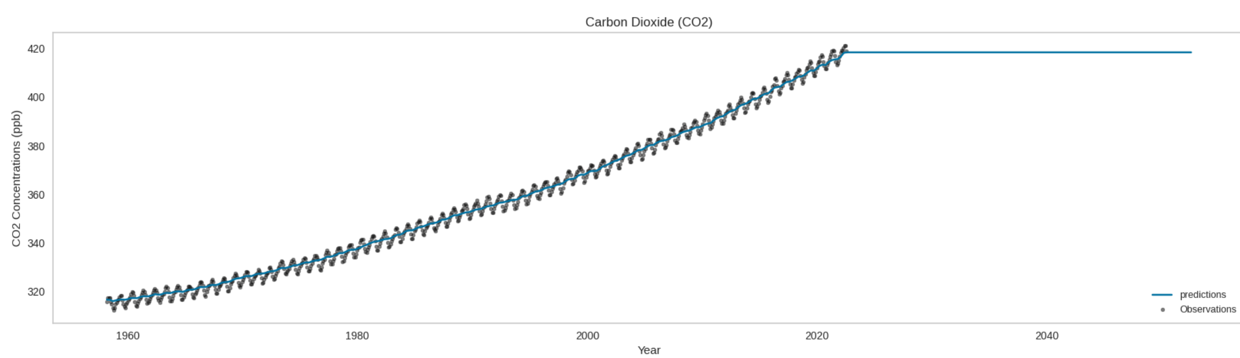
https://colab.research.google.com/drive/149krArphifM9NMV2JTtZBzkShf_Z6cqh?usp=sharing (Other GPR Models)

https://colab.research.google.com/drive/1wISpXnP65pAKA4AOhkS3XVctcEOrvMFM?usp=sharing (Other ML models)

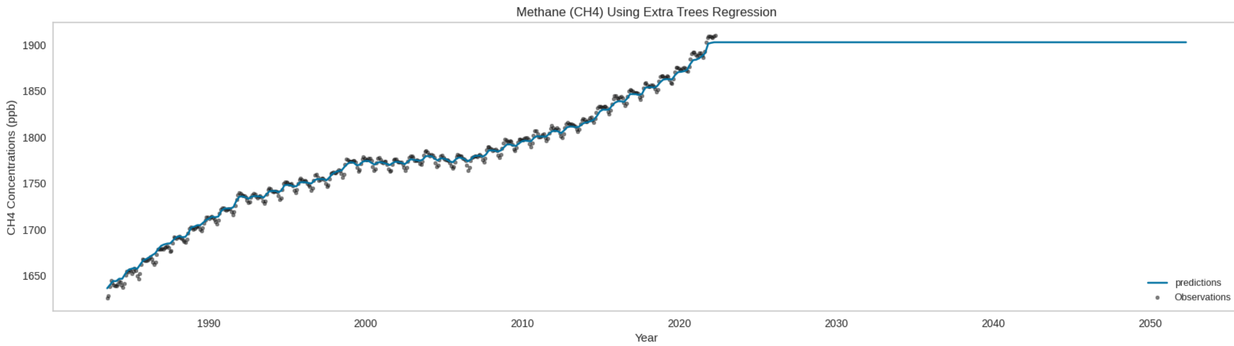
Appendix B (Images)



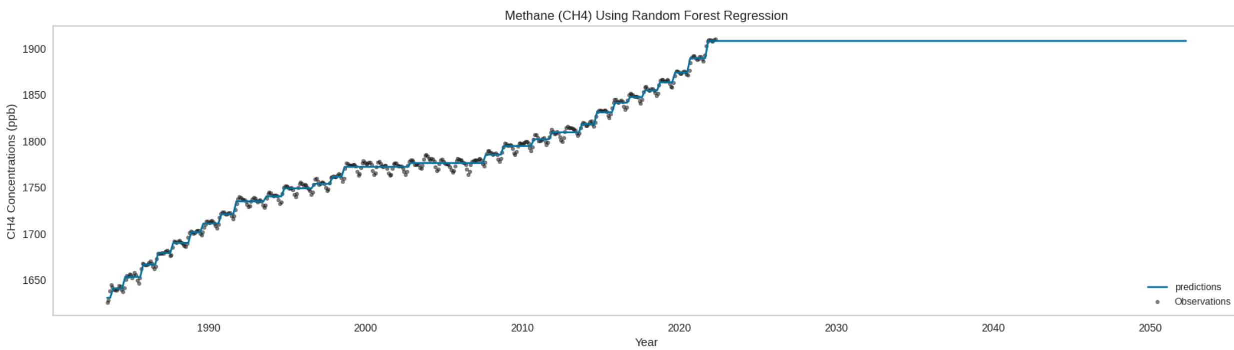
Appendix B1 CO2 Random Forest Model Plot



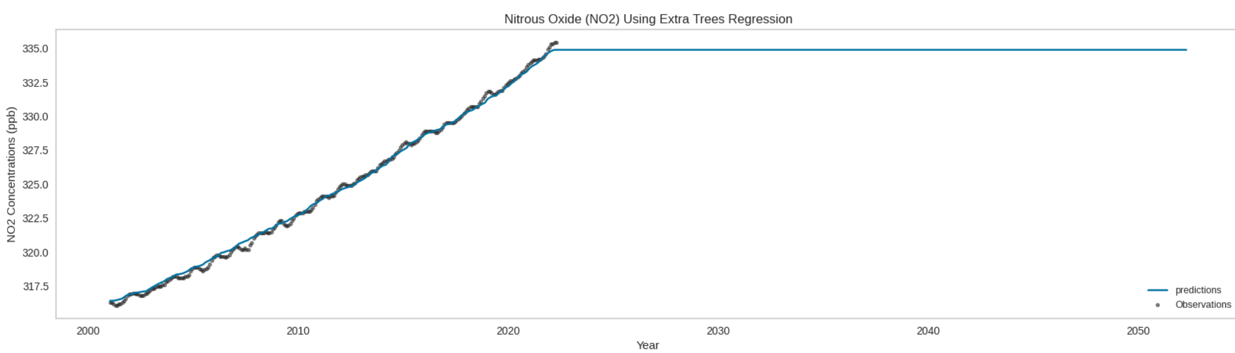
Appendix B2 CO2 Extra Trees Model Plot



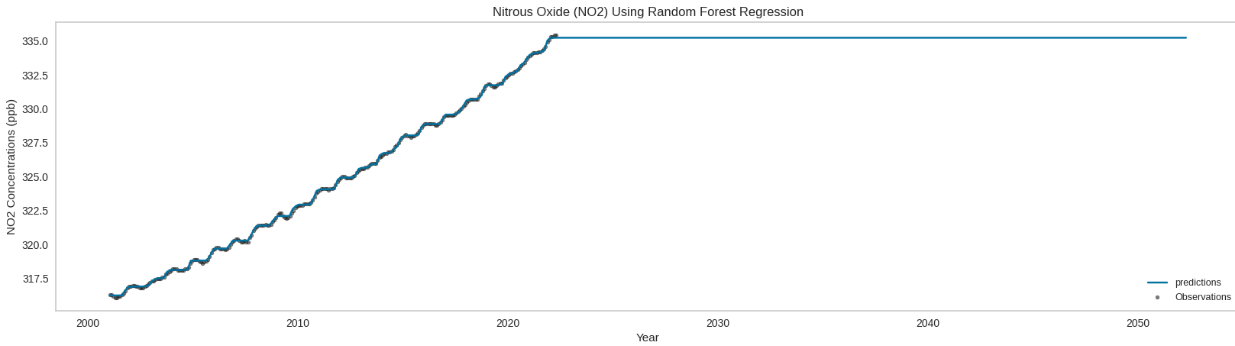
Appendix B3 CH4 Extra Trees Model



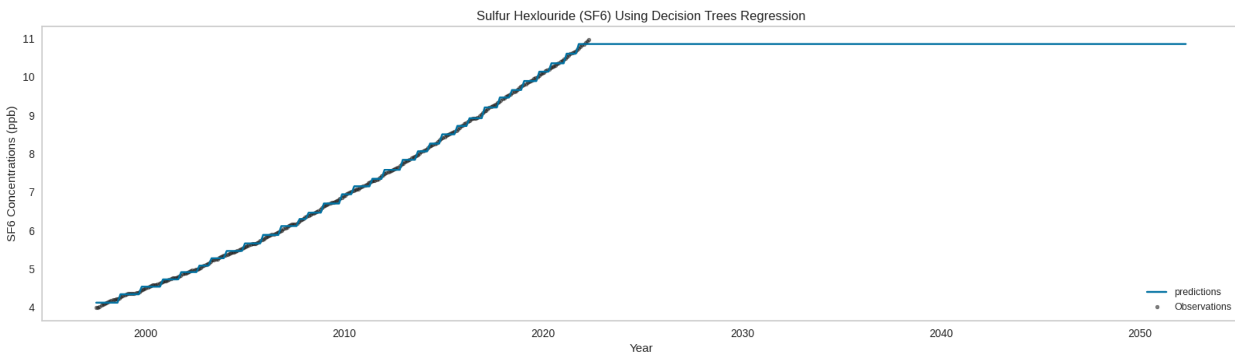
Appendix B4 CH4 Random Forest Model



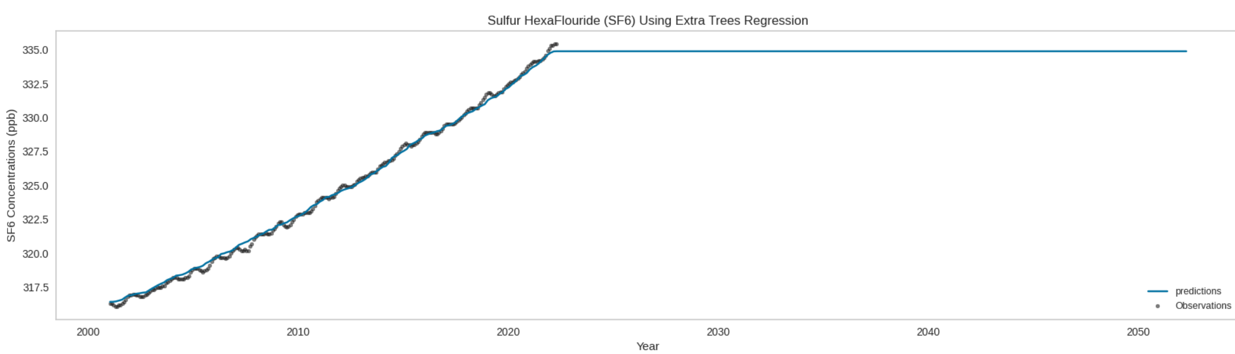
Appendix B5 N2O Extra Trees Model



Appendix B6 N2O Random Forest Model



Appendix B7 SF6 Decision Trees Model



Appendix B8 SF6 Extra Trees Model

Abstract

In this paper, a Fuzzy Logic Controller (FLC) is implemented for an Ambient Assisted Living/Enhanced Living Environment (AAL/ELE) System. The controller works by generating the appropriate control outputs from the inputs given to it. The controller was modeled as a Mamdani Fuzzy Inference System and was created in the MATLAB Fuzzy toolbox. After the implementation, the controller was then optimized using a Genetic Algorithm (GA) from the Global Optim Toolbox in MATLAB.

Keywords: **Ambient Assisted Living System, MATLAB, Genetic Algorithm, Mamdani Fuzzy Inference System, Optimization.**

1. Introduction

The world is experiencing a rapidly aging population, especially in developed countries. The population of elderly people in Europe is slightly above a quarter and is projected to increase even more (Li et al., 2015, 229-252). As elderly people are at higher risk of serious illness, motor handicaps, and cognitive decline, there is a need for a system that can provide solutions to improve the living conditions of the elderly, disabled, and patients living with chronic illnesses. This system should be user-friendly, robust, safe, and above all affordable.

An Ambient Assisted Living/Enhanced Living Environment (AAL/ELE) System is a system that supports the daily activities of elderly, disabled people, and patients living with chronic illnesses (Kara et al., 2017, 392-399). Such a system will have several requirements, a few of which are:

- They need to be able to adapt to the environment.
- They should be personalizable, i.e they should adapt to the user
- Should be non-invasive, and affordable.

Several models of such a system have been proposed and implemented. This report will focus on such a system built around a Fuzzy Logic Controller (FLC) and optimized using Genetic Algorithms.

The parameters that are controlled using the FLC in this report are: house temperature, patient blood pressure, patient location using a PIR sensor, heart rate, and light levels. These are believed sufficient to accurately and effectively model and maintain the patient’s wellbeing as well as help in keeping down energy consumption. The FLC, and its optimization using the Genetic Algorithm was achieved in MATLAB using the Fuzzy toolbox and the Global Optim Toolbox in MATLAB respectively..

2. System Design

The system in question is a Fuzzy Logic Control (FLC) System and as such is made up of several components. What an FLC does is map an input space to an output space using a set of fuzzy rules and linguistic values. These values are described by membership functions.

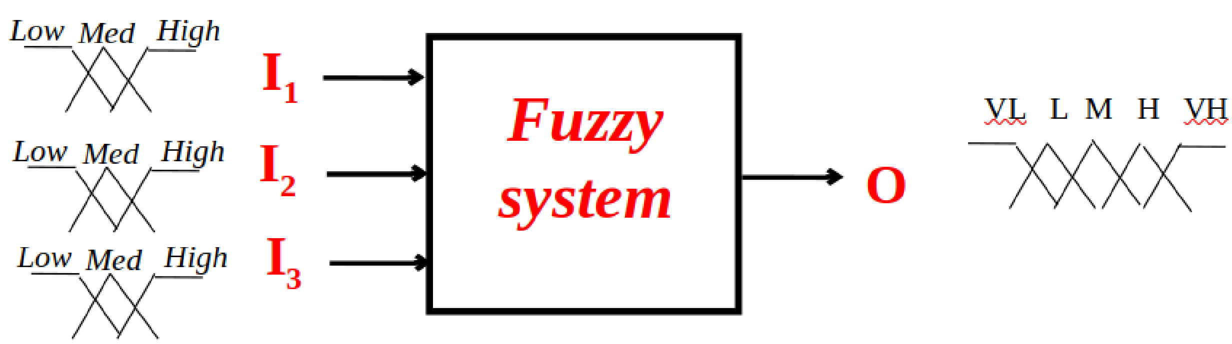


Figure 1: A high level overview of a fuzzy system (Doctor & Palade, 2022)

The rules for the fuzzy system are outlined in the following format: If I_1 is high and I_2 is low and I_3 is medium, then O is high. There main fuzzy models: Mamdani Fuzzy model and Sugeno Fuzzy model. This report makes use of a Mamadani model. The inner workings of fuzzy logic controller is shown below:

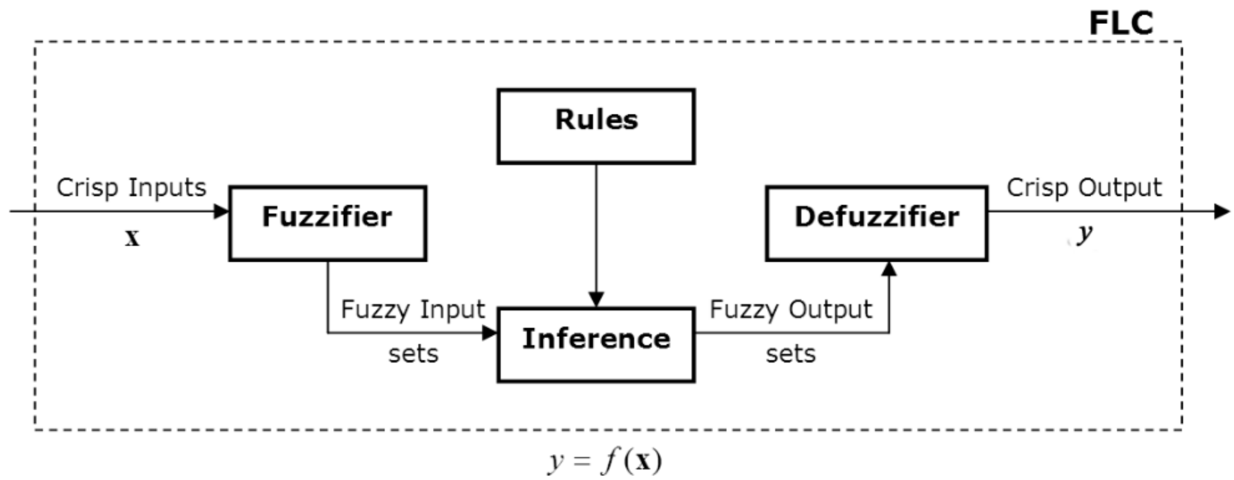


*Figure 2: An internal view of a Fuzzy Logic System (*Doctor & Palade, 2022*)*

The Mamdani Fuzzy Inference System model was proposed by Ebhasim Mamdani and makes use of linguistic rules to map input rules to output spaces. In this model, both the input and outputs are fuzzy sets. This means the output needs a membership function as well and requires defuzzification. It is more computationally intensive than its counterpart, the Sugeno Inference System, but is better suited to human input.

2.1 Inputs and Membership Functions

The FLC was designed to take five inputs. These are: house temperature, patient temperature, heart rate, light intensity, blood pressure, PIR sensor reading (for patient location). These parameters are considered sufficient to model and ensure the patient's wellbeing for the purpose of this report. The following section will show how the membership function of each variable/parameter was modeled.

2.1.1. House Temperature

The temperature of the house is one of the most important parameters. This parameter plays a significant role towards ensuring the patient's wellbeing and in extreme cases can even lead to the death of the patient. Hence, the temperature must be modeled accurately.

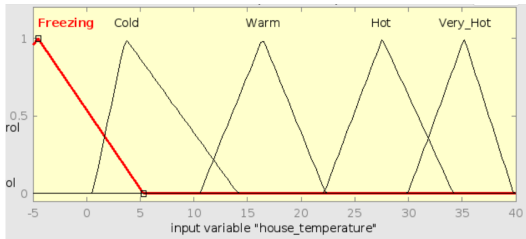


Figure 3: Temperature membership functions

Membership Function for Temperature is divided into Freezing, Cold, Warm, Hot and Very Hot while the range in Temperature within the house is taken from -5 to 40 degrees celsius.

Temperature Membership Function			
S/N	**Membership Functions**	**Range in degrees celsius**	**Type**
1	Freezing	-5 to 5°C	Triangle
2	Cold	0 to 14°C	Triangle
3	Warm	10 to 22°C	Triangle
4	Hot	22 to 34°C	Triangle
5	Very Hot	30 to 40°C	Triangle

Table 1: Temperature membership function

2.1.2 PIR Sensor

The PIR sensor is a proximity detector. It is used to detect if the patient is in the room or not. This can help with managing energy consumption as the lights can be turned off if the patient is not using them.

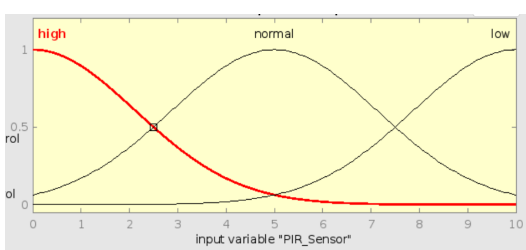


Figure 4: Proximity membership functions

Membership Function for PIR Sensor is divided into High, Normal and Low while the range in distance the PR Sensor detects is taken from 0 to 10 meters.

PIR Sensor Membership Function			
S/N	**Membership Functions**	**Range in metres**	**Type**
1	High	0 to 5m	Gaussian
2	Normal	0 to 10m	Gaussian
3	Low	5 to 10m	Gaussian

Table 2: PIR membership functions

2.1.3 Light Levels

This refers to the ambient lighting of the environment. If there is enough natural light, then there is no need for the artificial lighting to come on.

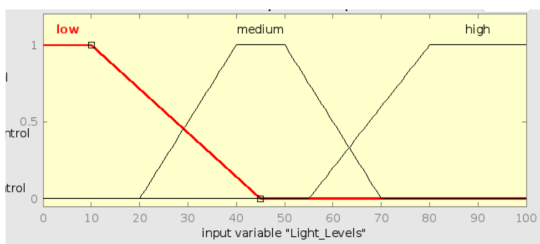


Figure 5: Membership function for light levels

Membership Function for Light Levels is divided into Low, Medium and High while the intensity of the Light Levels is taken from 0 to 100 thousand Lux.

Light Levels Membership Function			
S/N	**Membership Functions**	**Range in Lux**	**Type**
1	Low	0 to 45 Lux	Trapezoid
2	Medium	20 to 70 Lux	Trapezoid
3	High	55 to 100 Lux	Trapezoid

Table 3: Light levels membership functions

2.1.4 Blood Pressure

The blood pressure should be self-explanatory. The blood pressure is measured as

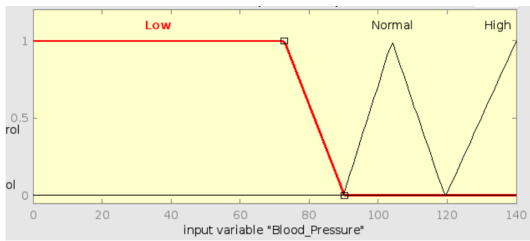


Figure 6: Blood pressure membership functions

The membership function values for Blood Pressure is divided into Low, Normal and High while the range of blood pressure is taken from 0 to 120 millimetres of mercury(mmHg)

Blood Pressure Membership Function			
S/N	**Membership Functions**	**Range in millimetres of mercury(mmHg)**	**Type**
1	Low	0 to 90mmHg	Trapezoid
2	Normal	90 to 120mmHg	Triangle
3	High	120 to 140mmHg	Trapezoid

Table 4: Blood pressure membership functions

2.1.5 Heart Rate

The heart rate could be used to model the mood or the state of agitation of the patient. Using this, the system can tell if the patient is anxious, or stressed.

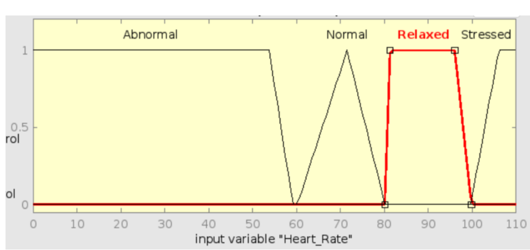


Figure 7: Heart rate membership functions

Membership Function for Heart Rate is divided into Abnormal, Normal, Relaxed and Stressed while the range in measurement of the heart rate is taken from 0 to 110 beats per minute(BPM).

Heart Rate Membership Function			
S/N	**Membership Functions**	**Range in beats per minute(BPM)**	**Type**
1	Abnormal	0 to 60bpm	Trapezoid
2	Normal	60 to 80bpm	Triangle
3	Relaxed	80 to 100bpm	Trapezoid
4	Stressed	100 to 110bpm	Trapezoid

Table 5: Heart rate membership functions

2.1.5 Body Temperature

The body temperature is one of the first symptoms that manifests when a patient is unwell. This checks if it is within the range of values specified below.

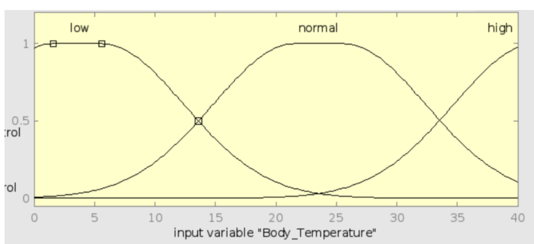


Figure 8: Body temperature membership functions

Membership Function for Body Temperature is divided into three; Low, Normal, and High while the measurement of the body temperature is within the range of 0 to 40 degrees celsius.

Body Temperature Membership Function			
S/N	**Membership Functions**	**Range in degrees celsius**	**Type**
1	Low	0 to 23°C	Gaussian
2	Normal	0 to 40°C	Gaussian
3	High	24 to 40°C	Gaussian

Table 6: Body temperature membership functions

2.2 Outputs

2.2.1 Light Control: LED lights, bulbs or lamps

This is the control of the artificial lighting of the apartment. It is mainly affected by the natural lighting and the proximity of the patient.

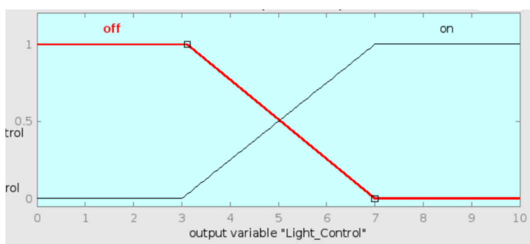


Figure 9: Light control membership functions

Light Control Membership Function			
S/N	**Membership Functions**	**Range in degrees celsius**	**Type**
1	Low	0 to 23°C	Trapezoid
2	Normal	0 to 40°C	Trapezoid

Table 7: Light control membership functions

2.2.2 Temperature Control:

This referrers to the general control of the heating and cooling appliances or systems

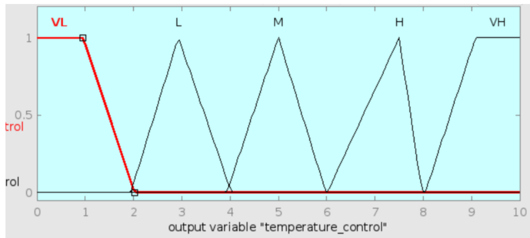


Figure 10: The temperature control membership functions

Temperature Control Membership Function			
S/N	**Membership Functions**	**Range in degrees celsius**	**Type**
1	Freezing	-5 to 5°C	Trapezoid
2	Cold	0 to 14°C	Triangle
3	Warm	10 to 22°C	Triangle
4	Hot	22 to 34°C	Triangle
5	Very Hot	30 to 40°C	Trapezoid

Table 8: Temperature control membership functions

2.2.3 Medication Control

This refers or determines if a resident in the disabled facility or care home will be administered drugs with the assistance of a nurse or not and if the situation requires the attention of paramedics or ambulance.

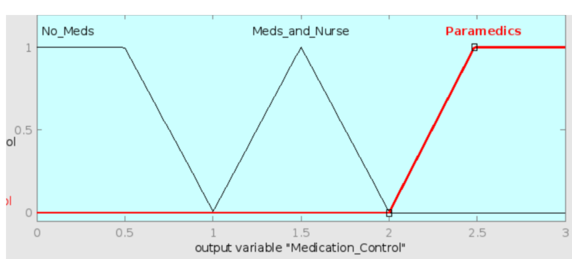


Figure 11: The membership functions of the medication control output

Medication Control Membership Function			
S/N	**Membership Functions**	**Range in degrees celsius**	**Type**
1	Low	0 to 23°C	Trapezoid
2	Normal	0 to 40°C	Triangle
3	High	24 to 40°C	Trapezoid

Table 9: Medication control membership functions

2.4 Fuzzy Rules

The fuzzy rules are usually generated by a domain expert. These rules tell us what the relationship between the inputs and the outputs are.

2.4.1 PIR Sensor and Light Levels for Light Control

PIR Sensor / **Light Levels**	**High**	**Normal**	**Low**
Low	ON	ON	OFF
Medium	ON	OFF	OFF
High	OFF	OFF	OFF

Table 10: Fuzzy rules for PIR sensor (input) and light controls (output)

2.4.2 Blood Pressure and Heart Rate vs Medication Control

Blood Pressure / **Heart Rate**	**High**	**Normal**	**Low**
Abnormal	Paramedics	Meds and Nurse	Paramedics
Normal	Meds and Nurse	No Meds	Meds and Nurse
Relaxed	Meds and Nurse	No Meds	Meds and Nurse
Stressed	Paramedics	Paramedics	Paramedics

Table 11: Fuzzy rules for blood pressure, heart rate (input), and medication control (output)

2.4.3 House Temperature and Blood Pressure vs Medication Control

Blood Pressure / **House Temperature**	**High**	**Normal**	**Low**
Freezing	Meds and Nurse	No Meds	Meds and Nurse
Cold	Paramedics	No Meds	Meds and Nurse
Warm	Meds and Nurse	No Meds	Meds and Nurse
Hot	Paramedics	No Meds	Meds and Nurse
Very Hot	Paramedics	No Meds	Meds and Nurse

Table 12: Rules for house temperature, blood pressure (input), and medication (output)

2.4.4 House Temperature and Light Levels vs Temperature Control

Body Temperature / **House Temperature**	**High**	**Normal**	**Low**
Freezing	Meds and Nurse	No Meds	Meds and Nurse
Cold	Paramedics	No Meds	Meds and Nurse
Warm	Meds and Nurse	No Meds	Meds and Nurse
Hot	Paramedics	No Meds	Meds and Nurse
Very Hot	Paramedics	No Meds	Meds and Nurse

Table 13: Rules for house temperature, light levels (input), and temperature controls (output)

The rest of the rules are too complex to be represented in table form. These rules have been added as snippets to Appendix A. The 3D plots of the output clippings have been added there as well (Appendix B).

3. Optimization Using Genetic Algorithm

Genetic Algorithms are a branch of evolutionary algorithms that were derived from studying natural phenomena. They are search and optimization algorithms that were introduced by John Holland in 1978. Using genetic algorithms, each solution to a problem is encoded as a sequence of bits called a chromosome. Each chromosome will then contain a number of bits called a gene that is represented by either a 0 or a 1 (Negnevitsky, 2011).

The algorithm works as follows: by applying the following major steps:

- Represent the problem variable domain as a chromosome of fixed length. Set the size of the population *N,* the crossover prob

3.1 Fitness Function

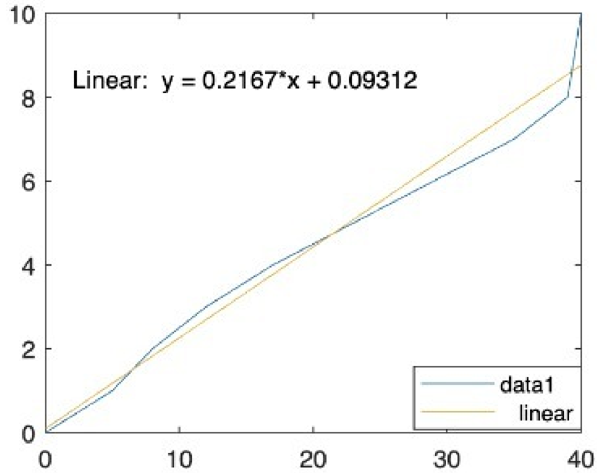
First iteration

Using the body temperature range from our fuzzy logic in task 2.1, 10 values within the range of 0 to 40 were chosen for variable x while random values from 0 to 10 were chosen from the temperature control scale of our fuzzy logic output to generate the values for y.

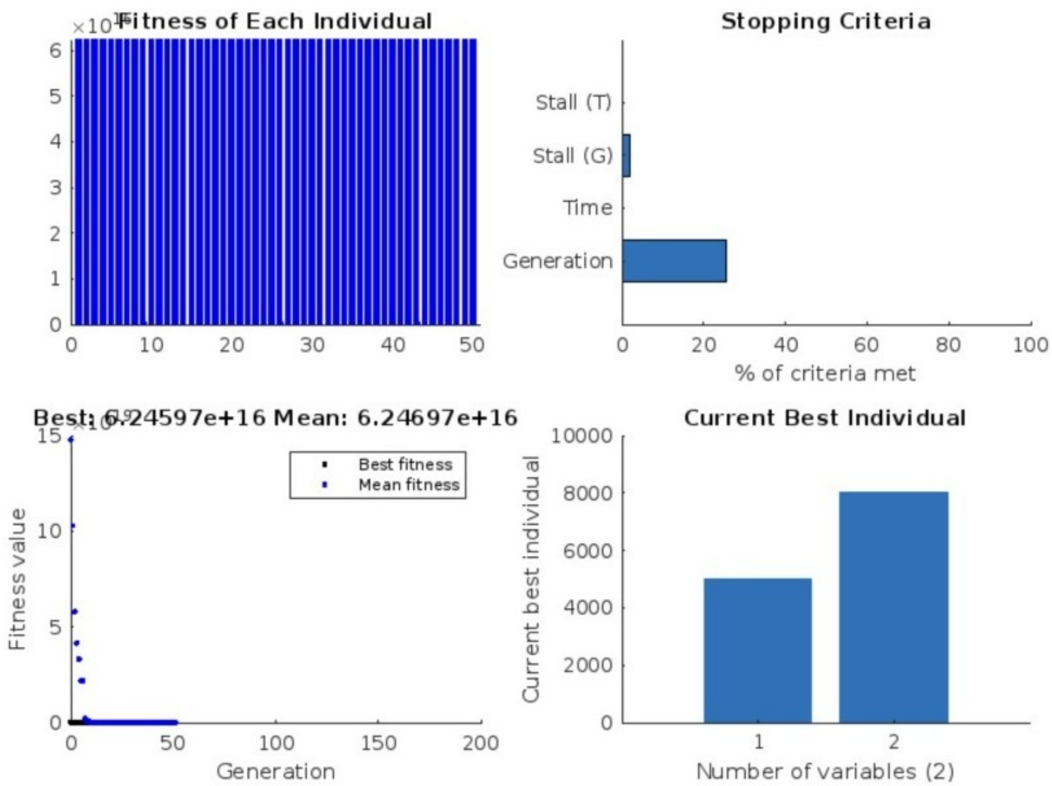
x = [0, 5, 8, 12, 17, 23, 29, 35, 39, 40]

y = [0, 1, 2, 3, 4, 5, 6, 7, 8, 10]

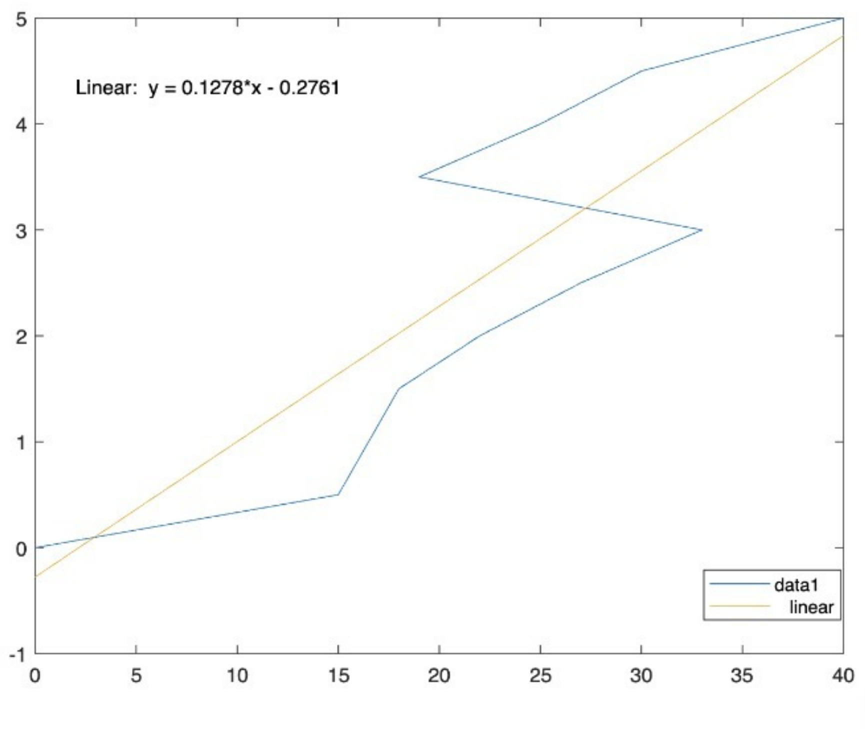
The variables x and y were plotted using the plot() function in MatLab which generated the fitness function below: y=0.2167*x(1) + 0.09312;

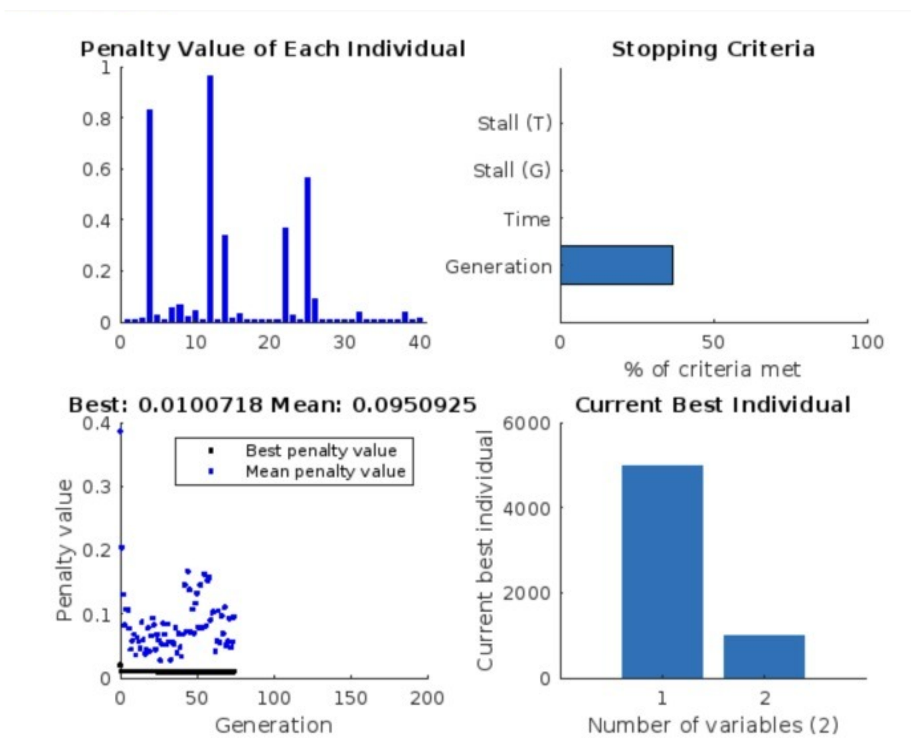


LB: Lower Boundary, UB: Upper Boundary
The LB and UB range were chosen randomly



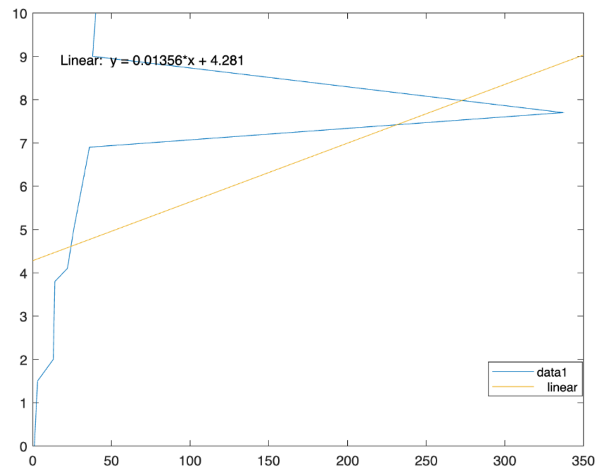
Second iteration
x = [0, 15, 18, 22, 27, 33, 19, 25, 30, 40]
y = [0, 0.5, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5]

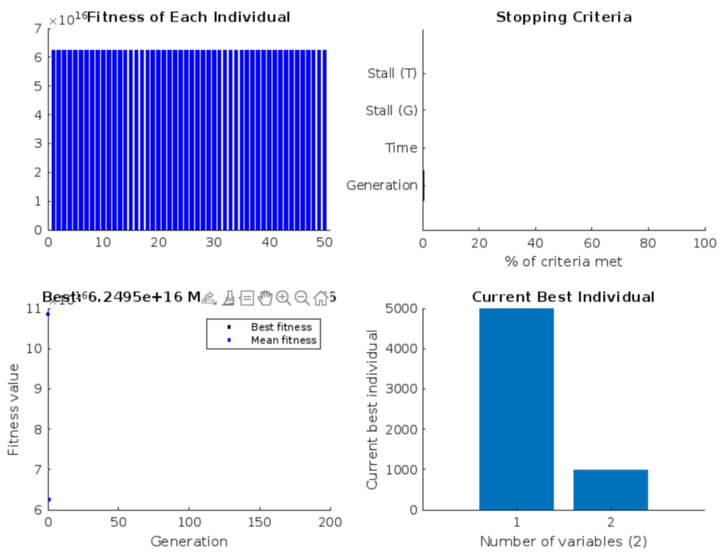




x = [1, 3, 13, 14, 22, 26, 36, 337, 38, 40]

y = [0, 1.5, 2, 3.8, 4.1, 5, 6.9, 7.7, 9, 10]


Linear: y = 0.01356*x + 4.281
data1
linear


Fitness of Each Individual
Stopping Criteria
Stall (T)
Stall (G)
Time
Generation
% of criteria met
Best: 6.2495e+16 M
Best fitness
Mean fitness
Fitness value
Generation
Current Best Individual
Current best individual
Number of variables (2)

4. Optimization Techniques on CEC's 2005 Functions

There are a variety of search and optimization algorithms. These include Genetic Algorithms, Particle Swarm Optimizations, etc. These algorithms are used to optimize parameters and by doing so, gain the best values from different functions. CEC provides a variety of functions on which these algorithms can be tested. In this section of the report, we run Genetic Algorithms and Particle Swarm Optimization on two of those functions, namely: Shifted Rotated Griewank's Function without Bounds (F7), and Shifted Rotated Ackley's Function with Global Optimum on Bounds (F8).

4.1 Shifted Rotated Griewank's Function without Bounds

The function is given shown in the image below:

$$F_7(\mathbf{x}) = \sum_{i=1}^{D} \frac{z_i^{\,2}}{4000} - \prod_{i=1}^{D} \cos\left(\frac{z_i}{\sqrt{i}}\right) + 1 + f\_bias_7 \ , \ \mathbf{z} = (\mathbf{x} - \mathbf{o}) * \mathbf{M} \ , \ \mathbf{x} = [x_1, x_2, \dots, x_D]$$

D: dimensions

$\mathbf{o} = [o_1, o_2, \dots, o_D]$: the shifted global optimum

M': linear transformation matrix, condition number=3

M =**M'**(1+0.3|N(0,1)|)

The table below is for the optimization using Genetic Algorithm

CEC Function = 7, D=10		CEC Function = 7, D=2
NO.	ga_main_val	ga_main_val
1	-179.8171	-179.9547
2	-179.7064	-179.9889
3	-179.1381	-179.9800
4	-179.5046	-179.9854
5	-179.4995	-179.9423
6	-179.7399	-179.9995
7	-179.8386	-179.9777
8	-179.2705	-179.9998
9	-179.5732	-179.9867
10	-179.8563	-179.9564
11	-179.4850	-179.9684
12	-179.4014	-179.9924
13	-179.5166	-179.9900
14	-179.6246	-179.9630
15	-179.3553	-179.9886
MAX VALUE	-179.1381	-179.9423
MIN VALUE	-179.8563	-179.9998
MEAN VALUE	-179.5551	-179.9782
STANDARD DEVIATION	0.2125	0.0174

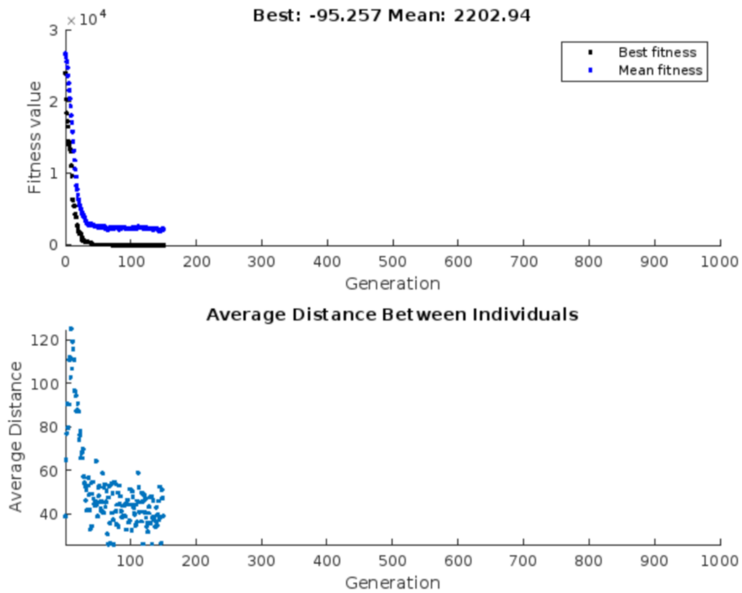


Figure 15: Graph of F7, d=10

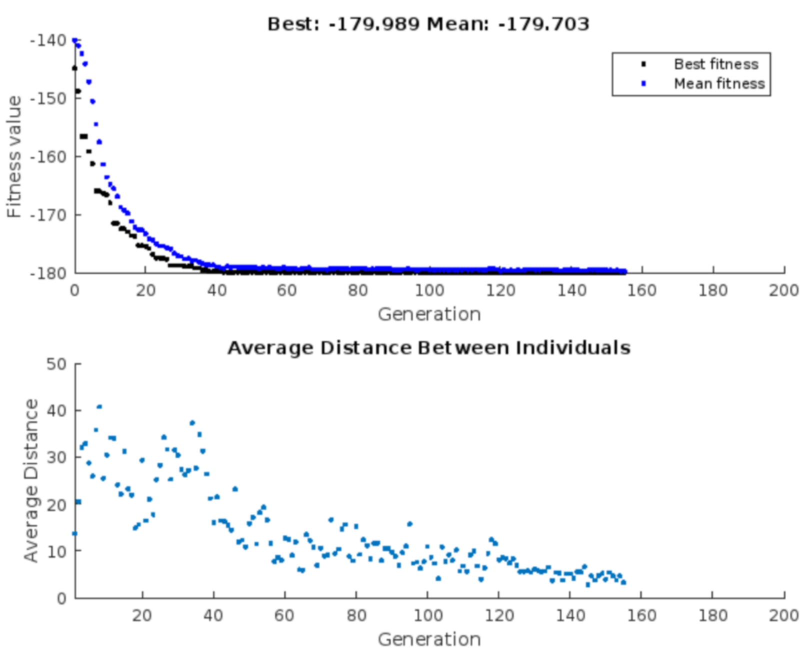


Figure 16: Graph of F7, d=2

4.2 Shifted Rotated Ackley's Function with Global Optimum on Bounds

$$F_8(\mathbf{x}) = -20\exp(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D} z_i^2}) - \exp(\frac{1}{D}\sum_{i=1}^{D}\cos(2\pi z_i)) + 20 + e + f\_bias_8,\ \mathbf{z} = (\mathbf{x}-\mathbf{o})*\mathbf{M},$$

$\mathbf{x} = [x_1, x_2, ..., x_D]$, D: dimensions

$\mathbf{o} = [o_1, o_2, ..., o_D]$: the shifted global optimum;

After load the data file, set $o_{2j-1} = -32$ o_{2j} are randomly distributed in the search range, for $j = 1, 2, ..., \lfloor D/2 \rfloor$

M: linear transformation matrix, condition number=100

The table below shows the above function details optimized using a genetic algorithm.

	CEC Function = 8, D=2	CEC Function = 8, D=10
NO.	ga_main_val	ga_main_val
1	-137.3352	-119.5228
2	-119.9841	-119.5354
3	-125.3657	-119.4652
4	-119.9689	-119.3968
5	-119.9919	-119.5324
6	-119.9903	-119.5101
7	-139.6951	-119.5287
8	-119.9845	-119.5693
9	-119.9980	-119.5281
10	-133.0738	-119.4910
11	-119.9948	-119.5536
12	-119.9995	-119.5758
13	-128.4362	-119.5011

14	-121.2840	-119.4411
15	-119.9966	-119.3596
MAX VALUE	-119.9689	-119.3596
MIN VALUE	-139.6951	-119.5758
MEAN VALUE	-124.3399	-119.5007
STANDARD DEVIATION	6.9612	0.0616

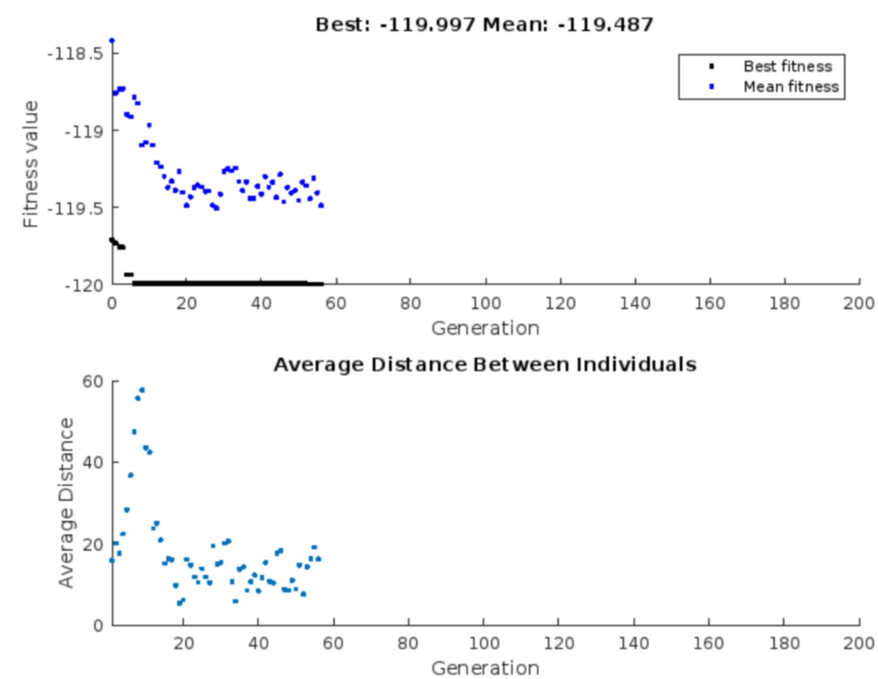


Figure 17: Function 8, D=2

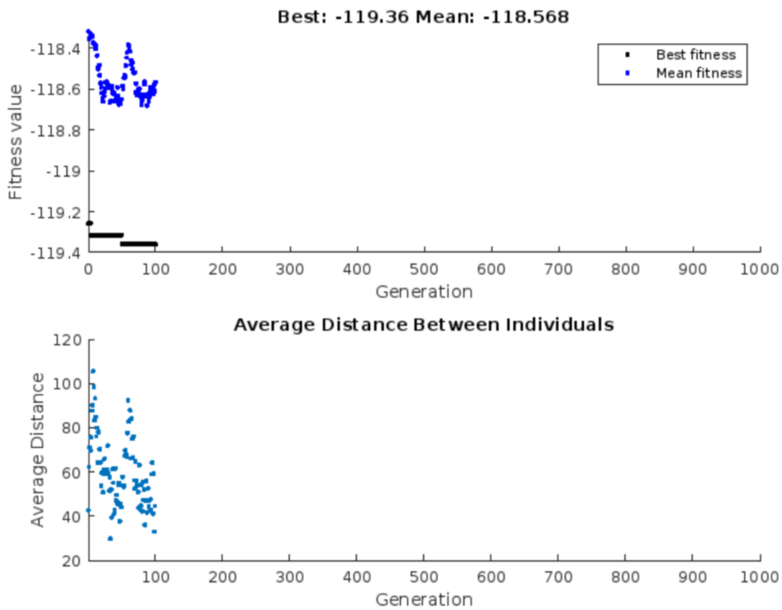
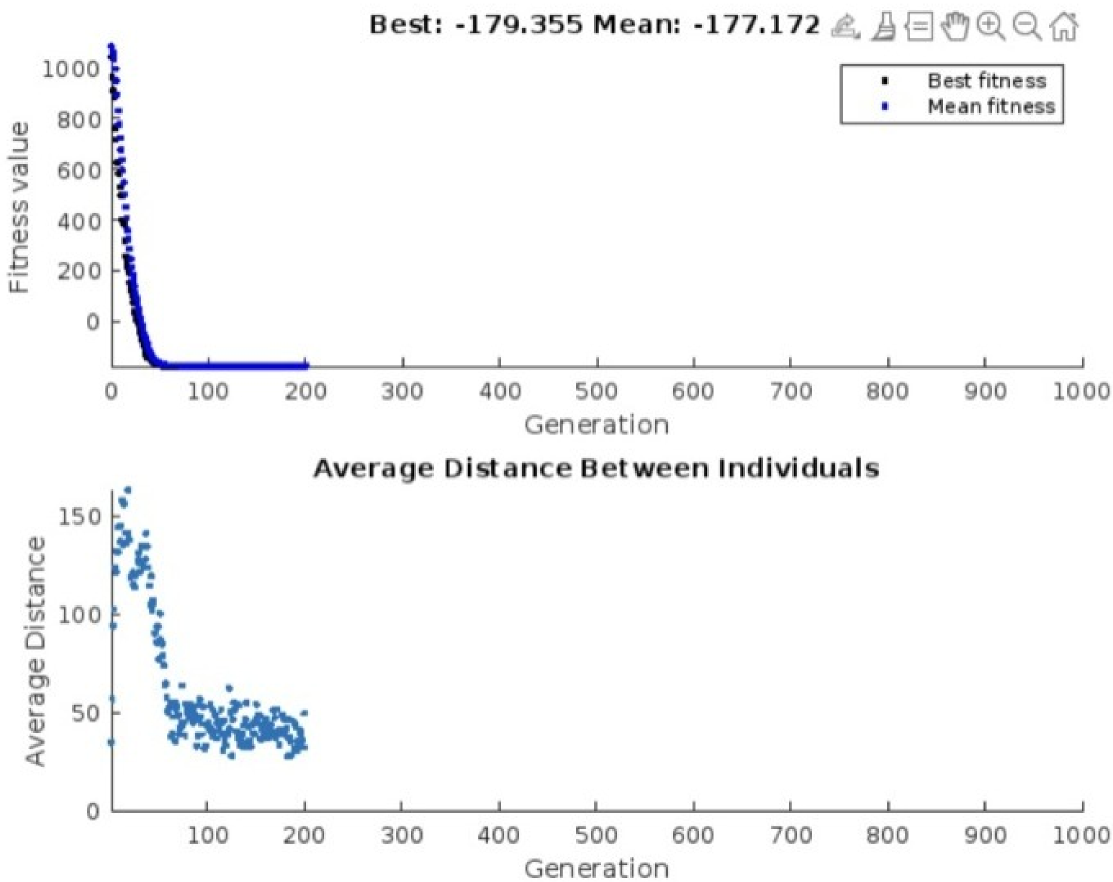
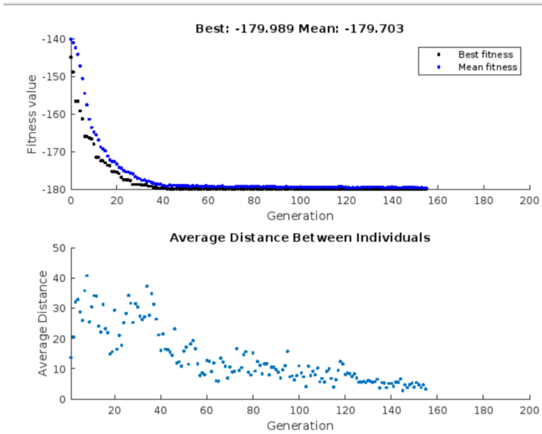


Figure 18: Function 8, D = 10


Best: -179.355 Mean: -177.172
Best fitness
Mean fitness
Fitness value
Generation
Average Distance Between Individuals
Average Distance
Generation


Best: -179.989 Mean: -179.703
Best fitness
Mean fitness
Fitness value
Generation
Average Distance Between Individuals
Average Distance
Generation

References

Doctor, F., & Palade, V. (2022). *Introduction to Fuzzy Logic 2* [Powepoint Presentation]. Coventry University, Coventry, United Kingdom. Retrieved August 17, 2022, from https://files.coventry.aula.education/bee415ea56fc03d3d97d17e4925f62a4lecture7.pptx

Kara, M., Lamouchi, O., & Ramdane-Cherif, A. (2017). A Quality Model for the Evaluation AAL Systems. *Procedia Computer Science*, *113*, 392-399. https://doi.org/10.1016/j.procs.2017.08.354

Li, R., Lu, B., & McDonald-Maier, K. D. (2015, November). Cognitive assisted Living Ambient System: A Survey. *Digital Communications and Networks*, *1*(4), 229-252. https://doi.org/10.1016/j.dcan.2015.10.003

Negnevitsky, M. (2011). *Artificial Intelligence: A Guide to Intelligent Systems* (3rd ed.). Pearson Education Limited.

Appendices

Appendix A (Rule Snippets)

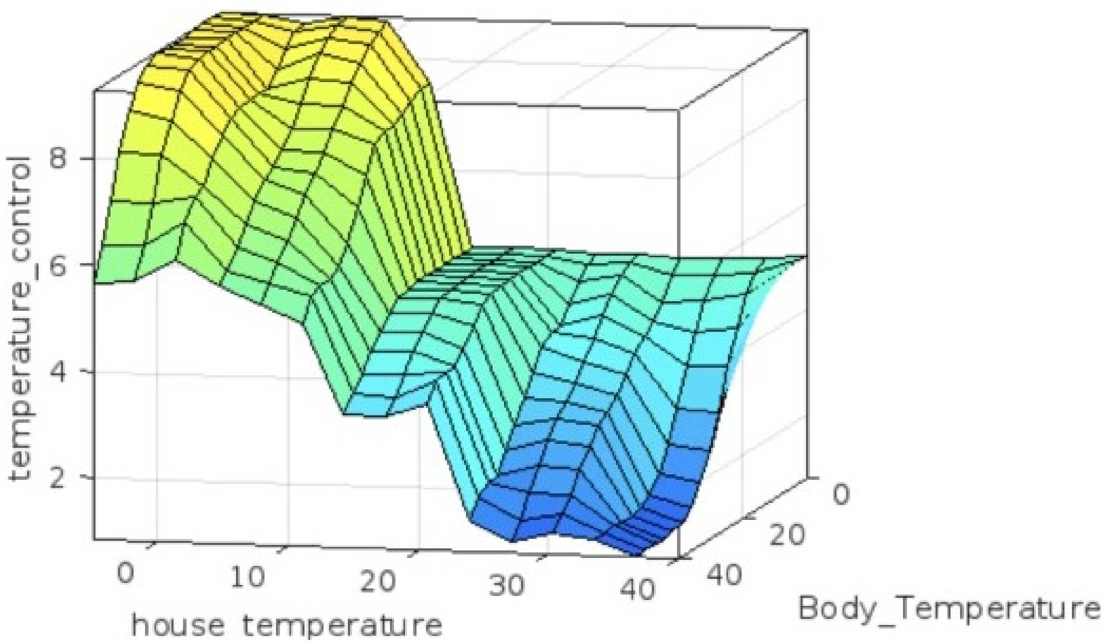
1. If (house_temperature is Freezing) and (Light_Levels is low) and (Body_Temperature is low) then (Light_Control is on)(temperature_control is VH) (1)
2. If (house_temperature is Cold) and (Light_Levels is low) and (Body_Temperature is low) then (Light_Control is on)(temperature_control is VH) (1)
3. If (house_temperature is Warm) and (Light_Levels is low) and (Body_Temperature is low) then (Light_Control is on)(temperature_control is H) (1)
4. If (house_temperature is Hot) and (Light_Levels is low) and (Body_Temperature is low) then (Light_Control is off)(temperature_control is M) (1)
5. If (house_temperature is Very_Hot) and (Light_Levels is low) and (Body_Temperature is low) then (Light_Control is off)(temperature_control is L) (1)
6. If (house_temperature is Very_Hot) and (Body_Temperature is normal) then (Light_Control is off)(temperature_control is VL) (1)
7. If (house_temperature is Hot) and (Body_Temperature is normal) then (Light_Control is off)(temperature_control is L) (1)
8. If (house_temperature is Warm) and (Body_Temperature is normal) then (Light_Control is off)(temperature_control is M) (1)
9. If (house_temperature is Cold) and (Body_Temperature is normal) then (Light_Control is off)(temperature_control is H) (1)
10. If (house_temperature is Freezing) and (Body_Temperature is normal) then (Light_Control is off)(temperature_control is VH) (1)
11. If (house_temperature is Freezing) and (Body_Temperature is high) then (Light_Control is off)(temperature_control is M) (1)
12. If (house_temperature is Cold) and (Body_Temperature is high) then (Light_Control is off)(temperature_control is M) (1)
13. If (house_temperature is Warm) and (Body_Temperature is high) then (Light_Control is off)(temperature_control is L) (1)
14. If (house_temperature is Hot) and (Body_Temperature is high) then (Light_Control is off)(temperature_control is VL) (1)
15. If (house_temperature is Very_Hot) and (Body_Temperature is high) then (Light_Control is off)(temperature_control is VL) (1)
16. If (house_temperature is Freezing) and (Body_Temperature is low) then (temperature_control is VH) (1)

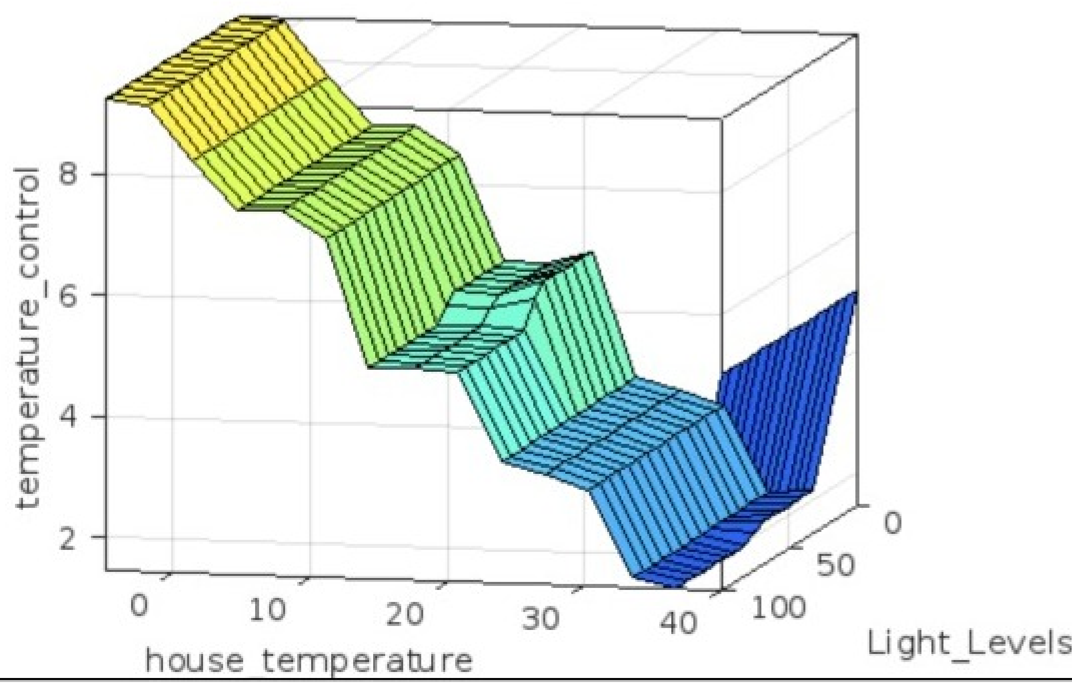
17. If (house_temperature is Cold) and (Body_Temperature is low) then (temperature_control is VH) (1)
18. If (house_temperature is Warm) and (Body_Temperature is low) then (temperature_control is M) (1)
19. If (house_temperature is Hot) and (Body_Temperature is low) then (temperature_control is M) (1)
20. If (house_temperature is Very_Hot) and (Body_Temperature is low) then (temperature_control is M) (1)
21. If (PIR_Sensor is high) and (Light_Levels is low) then (Light_Control is on) (1)
22. If (PIR_Sensor is high) and (Light_Levels is medium) then (Light_Control is on) (1)
23. If (PIR_Sensor is high) and (Light_Levels is high) then (Light_Control is off) (1)
24. If (PIR_Sensor is normal) and (Light_Levels is low) then (Light_Control is on) (1)
25. If (PIR_Sensor is normal) and (Light_Levels is medium) then (Light_Control is off) (1)
26. If (PIR_Sensor is normal) and (Light_Levels is high) then (Light_Control is off) (1)
27. If (PIR_Sensor is low) and (Light_Levels is low) then (Light_Control is off) (1)
28. If (PIR_Sensor is low) and (Light_Levels is medium) then (Light_Control is off) (1)
29. If (PIR_Sensor is low) and (Light_Levels is high) then (Light_Control is off) (1)
30. If (Blood_Pressure is Low) and (Heart_Rate is Relaxed) then (Medication_Control is Meds_and_Nurse) (1)
31. If (Blood_Pressure is Low) and (Heart_Rate is Abnormal) then (Medication_Control is Paramedics) (1)
32. If (Blood_Pressure is Low) and (Heart_Rate is Normal) then (Medication_Control is Meds_and_Nurse) (1)

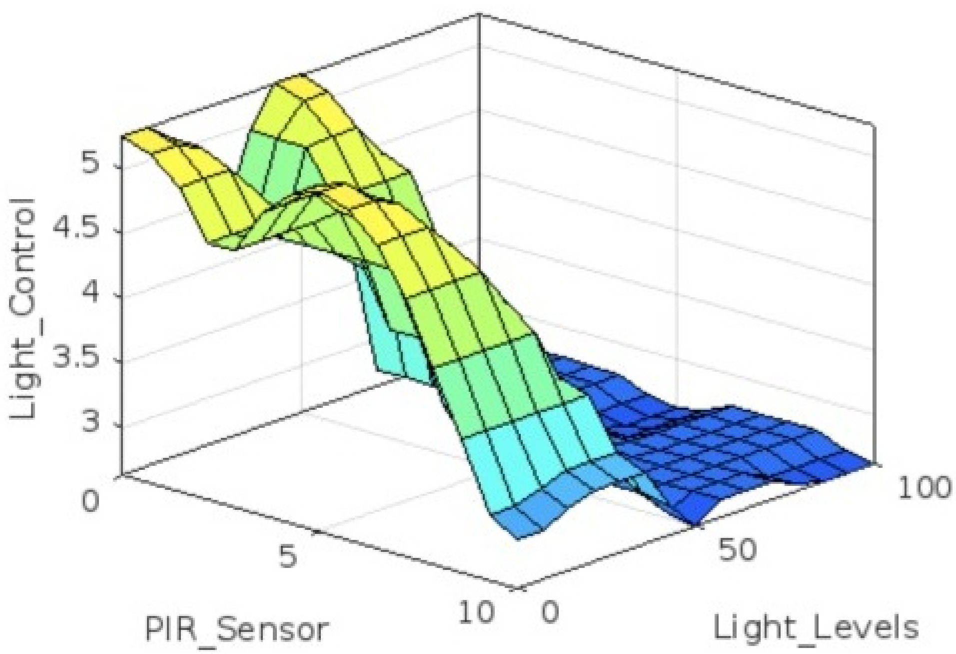
33. If (Blood_Pressure is Low) and (Heart_Rate is Stressed) then (Medication_Control is Paramedics) (1)
34. If (Blood_Pressure is Normal) and (Heart_Rate is Relaxed) then (Medication_Control is No_Meds) (1)
35. If (Blood_Pressure is Normal) and (Heart_Rate is Abnormal) then (Medication_Control is Meds_and_Nurse) (1)
36. If (Blood_Pressure is Normal) and (Heart_Rate is Normal) then (Medication_Control is No_Meds) (1)
37. If (Blood_Pressure is Normal) and (Heart_Rate is Stressed) then (Medication_Control is Paramedics) (1)
38. If (Blood_Pressure is High) and (Heart_Rate is Relaxed) then (Medication_Control is Meds_and_Nurse) (1)
39. If (Blood_Pressure is High) and (Heart_Rate is Abnormal) then (Medication_Control is Paramedics) (1)
40. If (Blood_Pressure is High) and (Heart_Rate is Normal) then (Medication_Control is Meds_and_Nurse) (1)
41. If (Blood_Pressure is High) and (Heart_Rate is Stressed) then (Medication_Control is Paramedics) (1)
42. If (house_temperature is Freezing) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
43. If (house_temperature is Freezing) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)
44. If (house_temperature is Freezing) and (Blood_Pressure is High) then (Medication_Control is Meds_and_Nurse) (1)
45. If (house_temperature is Cold) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
46. If (house_temperature is Cold) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)
47. If (house_temperature is Cold) and (Blood_Pressure is High) then (Medication_Control is Paramedics) (1)
48. If (house_temperature is Warm) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
49. If (house_temperature is Warm) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)

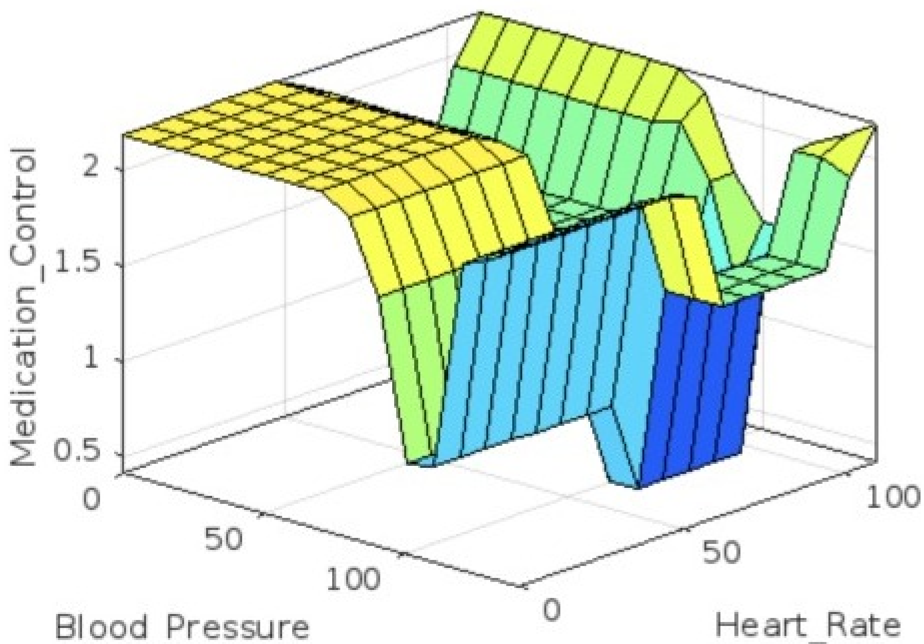
40. If (Blood_Pressure is High) and (Heart_Rate is Normal) then (Medication_Control is Meds_and_Nurse) (1)
41. If (Blood_Pressure is High) and (Heart_Rate is Stressed) then (Medication_Control is Paramedics) (1)
42. If (house_temperature is Freezing) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
43. If (house_temperature is Freezing) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)
44. If (house_temperature is Freezing) and (Blood_Pressure is High) then (Medication_Control is Meds_and_Nurse) (1)
45. If (house_temperature is Cold) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
46. If (house_temperature is Cold) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)
47. If (house_temperature is Cold) and (Blood_Pressure is High) then (Medication_Control is Paramedics) (1)
48. If (house_temperature is Warm) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
49. If (house_temperature is Warm) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)
50. If (house_temperature is Warm) and (Blood_Pressure is High) then (Medication_Control is Meds_and_Nurse) (1)
51. If (house_temperature is Hot) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
52. If (house_temperature is Hot) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)
53. If (house_temperature is Hot) and (Blood_Pressure is High) then (Medication_Control is Paramedics) (1)
54. If (house_temperature is Very_Hot) and (Blood_Pressure is Low) then (Medication_Control is Meds_and_Nurse) (1)
55. If (house_temperature is Very_Hot) and (Blood_Pressure is Normal) then (Medication_Control is No_Meds) (1)
56. If (house_temperature is Very_Hot) and (Blood_Pressure is High) then (Medication_Control is Paramedics) (1)

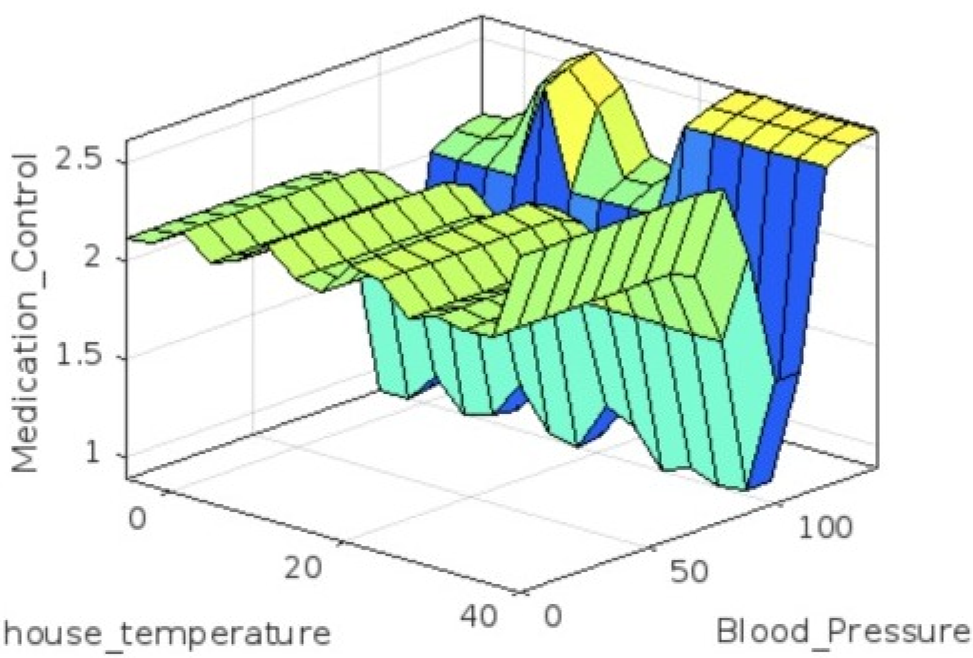
Appendix B (Clippings of the outputs)




temperature_control
8
6
4
2
0
10
20
30
40
house temperature
0
50
100
Light_Levels


Light_Control
5
4.5
4
3.5
3
0
5
10
PIR_Sensor
0
50
100
Light_Levels


Medication_Control
2
1.5
1
0.5
0
50
100
Blood_Pressure
0
50
100
Heart_Rate


Medication_Control
2.5
2
1.5
1
0
20
40
house_temperature
0
50
100
Blood_Pressure