Description of Advanced Features

All advanced features were implemented as part of developing the aggregation algorithm for light and temperature sensors, however, this document will only discuss the implementation details and challenges of the linear regression analysis of two vectors.

Once twelve new readings were obtained and filled the light and temperature sensor buffers, linear regression analysis was able to take place. It is necessary to wait for both buffers to be of the same size so that the regression line can be drawn between the same amount of points in the x and y-axis. In this implementation the x variable was the intensity of light measured in lux and the y variable was the temperature in degrees celsius. The gradient of the regression line can be found as it must pass through the mean points of the light and the temperature values in their respective buffers. This was calculated using the equation below.

$$gradient = \frac{\sum_{n=0}^{n-1} (x_n - x_{mean}) * (y_n - y_{mean})}{\sum_{n=0}^{n-1} (x_n - x_{mean})^2}$$

The intercept of the regression line can then be found by re-arranging $y = mx + c$ to $c = y - mx$. The values of y and x can be replaced by the mean values of the light and temperature readings in the buffer, as it is known that the regression line must pass through this point (x_{mean}, y_{mean}). This then provides the equation for the linear regression line.

$$y_{predicted} = gradient * x + intercept$$

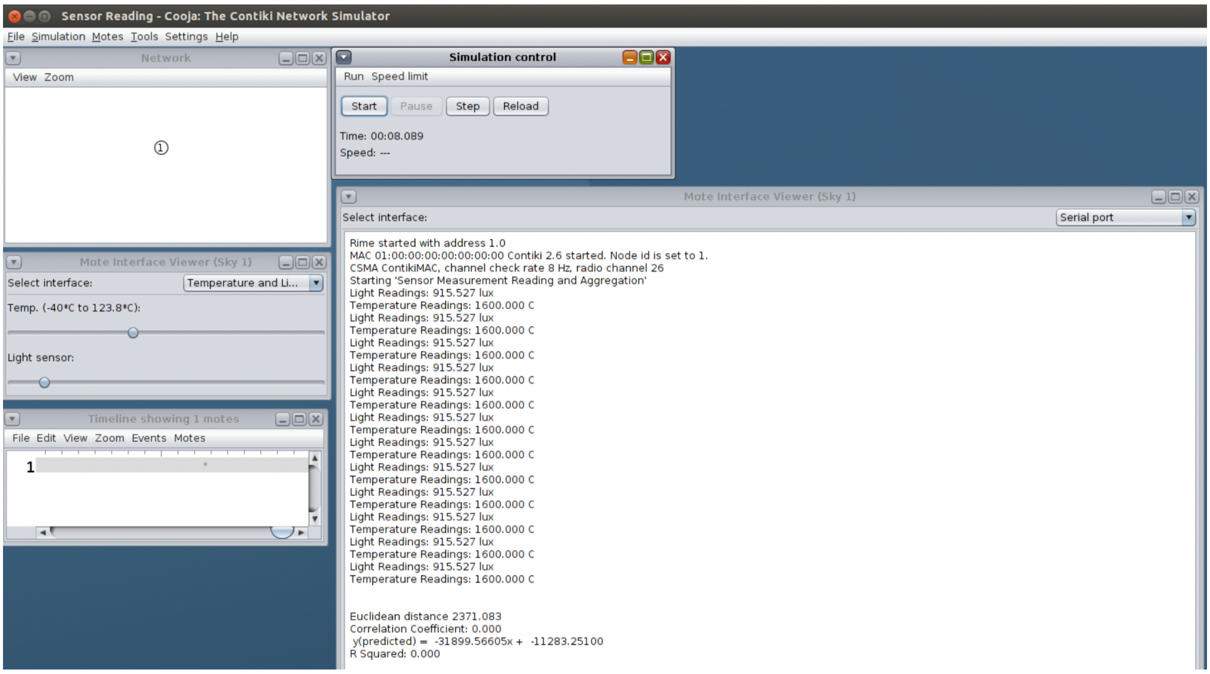
In order to evaluate how far the predicted y values are from the real values, R-Squared can be calculated. This can be achieved by first calculating the predicted temperature values by substituting the real light values for x in the linear regression equation, which can then be used in the following equation. Note: the highest value of R-Squared is 1 indicating that all predicted values are found exactly on the regression line.

$$R^2 = \frac{\sum_{n=0}^{n-1} (y_{pred} - y_{mean})^2}{\sum_{n=0}^{n-1} (y_n - y_{mean})^2}$$

The main challenge of implementing the linear regression analysis was programmatically trying to demonstrate an R-squared value that is 0 when the denominator of the equation was equal to 0. This was difficult as it is a challenge to be able to compare floats to be exactly 0; in order to get around this problem, the R-squared value was set to 0 if the denominator was not greater than 0 otherwise it would be calculated as seen above.

Screenshots demonstrating the linear regression analysis can be found on the next page.

Screenshot of Advanced Features (R-Squared at 0)



Screenshot of Advanced Features (R-Squared > 0)