5. Data Analysis I THINK WE NEED A WHOLE CHAPTER FOR THIS AS THERE ARE TOO MANY SUBSUBSUBSUBTITLES

Code for the below described data analysis methods can be found in appendix REFERENCE. Results foreach membrane colour and particle size were normalised by subtracting the control temperature at each time from nanoparticle membrane temperature. The resulting temperatures were then plotted as a function of time for each concentration measured. Uncertainties in individual temperatures were taken to be ±2°C from the thermal camera manual, and uncertainties in the average temperature values were found by the standard deviations of the data points given by the numpy function numpy.std and the formula (5.1):sn=√√√√1n−1n∑i=1(xi− ̄xn)2(5.1)where n is the number of data points present, xi are the individual mean temperature data points and xn is the mean of the data points.

We first investigated the change in nanoparticle membrane temperature with time relative to the concentration of nanoparticles on the membrane. It was assumed that the correlation between nanoparticle membrane temperature and time was linear and so least squares line of best fit was plotted using the scipy.stats.linregress function which utilises the following formulae to find intercept (c) and gradient (m)

((5.2) and19

(5.3)):c=∑iwixi2∑iwiyi−∑iwixi∑iwixiyi∆′(5.2)m=∑iwi∑iwixiyi−∑iwixi∑iwiyi∆′(5.3)With uncertainties in interceptαcand gradient being given by:αc=√∑iwixi2∆′(5.4)αm=√∑iwi∆′(5.5)Where∆′=∑iwi∑iwixi2−(∑iwixi)andxiandyiare the differential temperatures and time data points respectively.

SUBSUBTITLE (ONE WITHOUT NUMBERS): IRON OXIDE NANOPARTICLES

IRON OXIDE GRAPHS

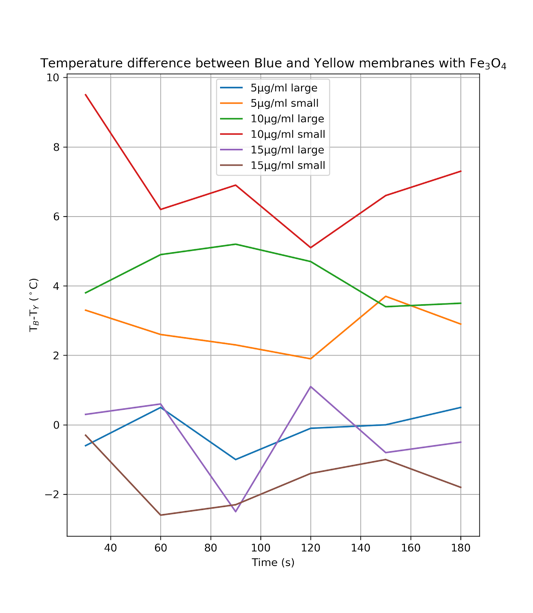
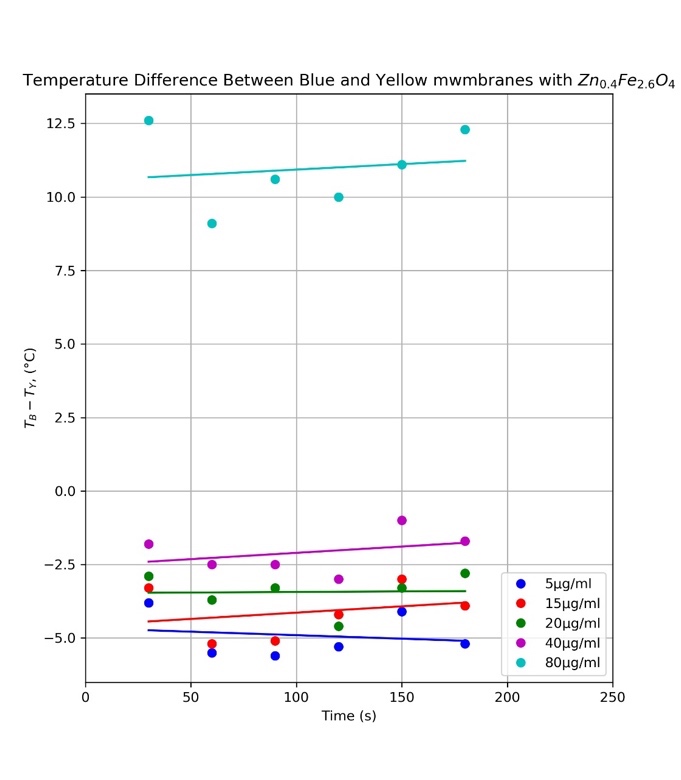
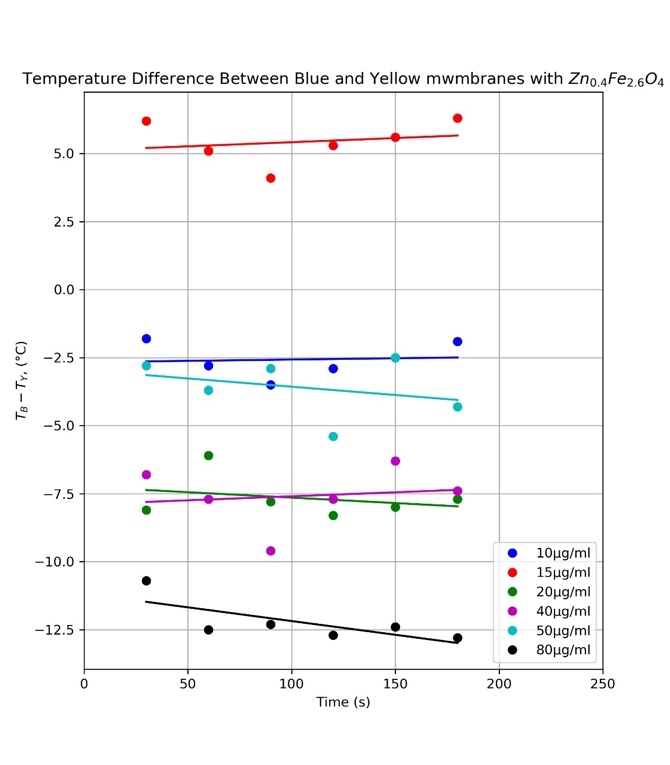
In theory, a higher concentration of nanoparticles should cause a greater increase in temperature when exposed to a laser [8]. However, this only holds true for graph in Figure 5.3a. For Figure 5.3b this assumption almost completely breaks down, with only the lower two concentrations being correctly positioned. Most significantly the 50μg/ml specimen has a mean temperature far lower than expected, at almost 20°Cbelow the average temperature for the40μg/ml concentration. Figure 5.3c shows a similar trend, with our 15μg/ml average temperature being located at nearly half the temperature of the 10μg/ml result. Figure 5.3d, is close to being as expected, with only the 50μg/ml concentration being placed strangely according to prior knowledge. It is a possible conclusion at this stage that perhaps the concentrations were made incorrectly, and they are not in fact the concentrations expected. The linear trends of the below graphs most frequently show a constant differential temperature with respect to time. Deviations from this seem to be random, and most likely due to random uncertainties as both positive and negative correlation changes can be seen with equal frequency.

SUBSUBTITLE (ONE WITHOUT NUMBERS): ZINC FERRITE NANOPARTICLES

ZINC FERRITE STRAIGHT LINE GRAPHS

Initially it was expected that differential temperature curve with time would undergo a rapid increase and then become more shallow with time. However, the experimental data does not show any increase in temperature, instead the temperature readings are fluctuating. Since the data measurements were taken with 30 s intervals, the rapid increase in temperature is therefore expected to happen during the first 30 s of the exposure to the laser. Results for Xb and Xd show the expected increase in mean temperature with concentration here. Xc shows an almost correct increase in mean temperature with concentration increase, with the significant outlier being the 80µg/ml concentration which shows a temperature roughly around the mean temperature of the 20µg/ml zinc ferrite concentration. This concentration should be treated as an anomaly and removed. Xa shows a confusing trend similar to that seen for some of the graphs of Iron Oxide nanoparticle changes in mean temperature with time, in that a few concentrations seem to be randomly ordered and not showing a trend. Similarly, it could be suggested that the concentrations here are incorrectly assumed.

SUBSUBTITLE Membrane Colour Investigation

To investigate the effect of different coloured membranes on our nanoparticle temperatures, graphs of results of mean temperature for the nanoparticles deposited onto the blue membranes minus the temperature for yellow membranes were plotted against time for both nanoparticle types:  
  


*Fig. X. I*

Iron Oxide nanoparticle results shown in Xa seem to imply mixed conclusions. We seem to see 3 trends of the mean temperature for the blue membrane being higher (10µg/ml small, 10µg/ml large and 5µg/ml small), 2 trends of the mean temperature for the yellow membrane being roughly around zero (5µg/ml large and 15µg/ml large) - albeit with large fluctuations for the 15µg/ml large result, and one result showing a higher temperature for the yellow membrane (15µg/ml small). There seems to be no correlation between large and small particles and concentrations from this data, implying that the nanoparticle size or concentration are not dependant on the membrane colour, but this fluctuation is stochastic.

Xb shows perhaps one of the most significant trends yet, with a clear larger temperature being seen for all concentrations deposited onto a yellow membrane. This data set also shows a significant outlier at 80µg/ml for unclear reasons.

Xc shows again an extremely significant trend, and one which is the same as what is shown in Xb. These results show clearly that the blue membrane absorbs more energy than the yellow one, which is in contradiction to previous results, although said previous results are arguably weaker in possibility.

5.1: The Limit of Detection

The limit of detection is defined as the lowest concentration that can be detected due to a temperature change of the nanoparticles on the membrane. In order to find our limit of detection we must plot graphs of mean temperature change of each nanoparticle relative to its concentration, plot a line of best fit and extrapolate this line of best fit to where there is zero mean temperature change. However the type of correlation of these data points was unclear, and therefore three types of lines of best fit were analysed, with results discussed in the subsections below.

Since the uncertainty of the individual temperature readings is±2°C, all of the readings that are less or equal to 2°C cannot be used to determine whether there are nanoparticles in the solution. The limit of detection is therefore determined by the resolution of the thermal camera.

The uncertainties in the concentration of nanoparticles in the solution were found by assuming two situations: when there is the maximum possible amount of nanoparticle concentrate with the minimum possible amount of water for the upper boundary and vice versa for the bottom boundary. The absolute uncertainty was found to be±0.5μg/ml. The percentage un-certainty is high for 5μg/ml solution, as it reaches 10%, however, as the concentration increases, the percentage uncertainty becomes < 1%and no longer significantly affects the results.

The uncertainties on the lines of best fit were estimated via a chi-squared fit of the curves relative to the data points with the Python function scipy.stats.chisquare which utilises the formula\cite{ziegel\_ott\_1994}:

INSERT CHI SQUARED FORMULA

WITH DEFINED FORMULA

This formula predicts the goodness of fit of a line to a set of data points, and therefore was intended for use as deciding which fit is most accurate to our trends, as well as giving a rough idea of uncertainty to lines for which uncertainties could not be calculated.

Reduced-chi squared shows us more clearly how good the model fit is, as it standardises the data. A $chisquared with a squiggle on top$ $much greater than sign the double one$ 1 indicates a poor model fit. A $chisquared with a squiggle on top$ $ greater than sign$ indicates that the fit has not fully captured the data, or that error variance has been underestimated. A $chi square = 1$ tells us that the observations and error estimations are in agreement and therefore the fit is ideal, and a $xhisquarelessthan1$ indicated an overfitted model; one which has either fitted noise as if it was a trend shown by the model itself or where the errors have been overestimated.

We also calculate the chi squared probability, which is also outputted from the above-mentioned Python function. This probability tells us how likely the fit is to be significant, with this particular function outputting a low value indicating that the fit is likely to be bad, and a high value indicating a good fit. A chi-squared probability greater than 0.95 is said to be extremely statistically significant, with there being a strong relationship between the models of our lines of best fit and our experimental data points. Please note that usually the significance is opposite for high and low p values, but in this case the function outputs one minus the probability.

5.1.1: A Linear Trend

It was initially assumed that the correlation between nanoparticle membrane temperature and con-centration was linear and so least squares line of best fit was plotted using the scipy.stats.linregress function which utilises the following formulae to find intercept (c) and gradient (m)

((5.2) and19

(5.3)):c=∑iwixi2∑iwiyi−∑iwixi∑iwixiyi∆′(5.2)m=∑iwi∑iwixiyi−∑iwixi∑iwiyi∆′(5.3)

With uncertainties in intercept αc and gradient being given by:αc=√∑iwixi2∆′(5.4)αm=√∑iwi∆′(5.5)Where∆′=∑iwi∑iwixi2−(∑iwixi)andxiandyiare the differential temperatures and time data points respectively.

SUBSUBTITLE (ONE WITHOUT NUMBERS): IRON OXIDE NANOPARTICLES

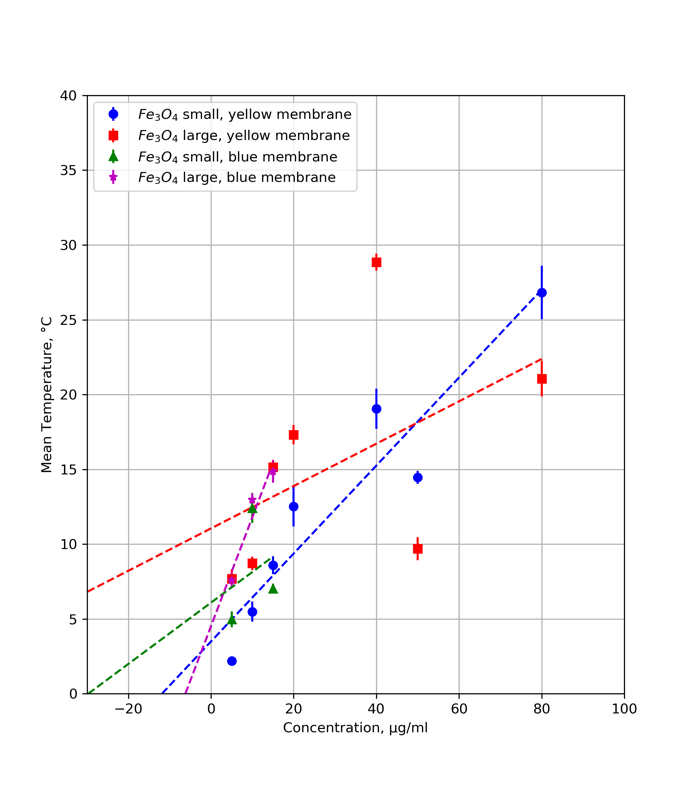


Fig X: *A graph of mean temperature reached by varying sizes of Iron Oxide nanoparticles on different coloured membranes against the concentration of nanoparticles present. Please note the lines of best fit are shown as extrapolated to the left past the data points to illustrate the limit of detection – the concentration at which the mean temperature tends to zero.*

This graph shows a significant difference between variation in mean temperature with increasing concentration when either blue or yellow membranes are used. For the small sized iron oxide nanoparticles we seem to see a similar gradient present for both yellow and blue membranes, but a roughly 5℃ increase in mean temperature of the blue membrane line for the same concentrations. This suggests that the yellow membrane could be absorbing far more of the laser energy. For the large iron oxide particles, we can see a radically different gradient. Mean temperature seems to increase far more dramatically per unit concentration for the blue membrane. It is possible that the large iron oxide particles on the blue membrane is an outlier due to this radically different gradient. We can see a difference in mean temperature of the lines again, except in this case our blue membrane reading is lower. Therefore whether blue or yellow membranes absorb more is inconclusive.

A table summarizing the limits of detection and chi squared fits can be seen in Table X below. We can see that all of our concentration limits are negative, rendering them unphysical and therefore inconclusive on what the limit of detection for these nanoparticles are.

Our chi squared values are predictably consistently low, with similarly low chi-squared probabilities and high reduced chi squared values indicating that our lines of best fit do not have a good relationship with our data points. However, there is one value that does not comply with this trend – the value for our large Iron Oxide nanoparticles deposited onto a blue membrane, which has a high chi-square probability indicating close to a good fit. However, the reduced chi square value for this estimation is extremely low, indicating that the model is in fact overfitted. This is visually the fit with the least outliers, the one that seems to fit closest to our data points, and the one which is least negative, so this output is predictable. We also have another good result, with our small yellow data showing a reduced chi squared value of exactly one, indicating a good fit. However, the probability is not good and the concentration is still non-physical. We must remember that our chi-squared probability only assesses how good the fit is to our data points, and not whether it fits with our experimental knowledge, and since all concentration values lie below zero for this fit, we can assume that either a linear fit is not correct, or our experimental data is flawed. We will conclude on this point after analysis of other fits.

It should be noted that since a majority of the lines of best fit of these data points lie outside their uncertainties, and therefore the results cannot really be said to be accurate or conclusive. Certainly, the uncertainties on concentration and mean temperature of the data points need to be reconsidered to be far greater.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |
|  | Small | | Large | |
| Membrane Colour | Yellow | Blue | Yellow | Blue |
| Limit Of Detection / μg/ml | -12+-7 | -30+/-114 | -78+-69 | -6+-3 |
| Chi Square | 4.5 | 3.4 | 16.5 | 0.2 |
| Reduced Chi Square | 0.8 | 1.7 | 2.7 | 0.09 |
| Chi-Square Probability | 0.6 | 0.2 | 0.01 | 0.9 |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

SUBSUBTITLE NO NUMBERS: ZINC FERRITE NANOPARTICLES

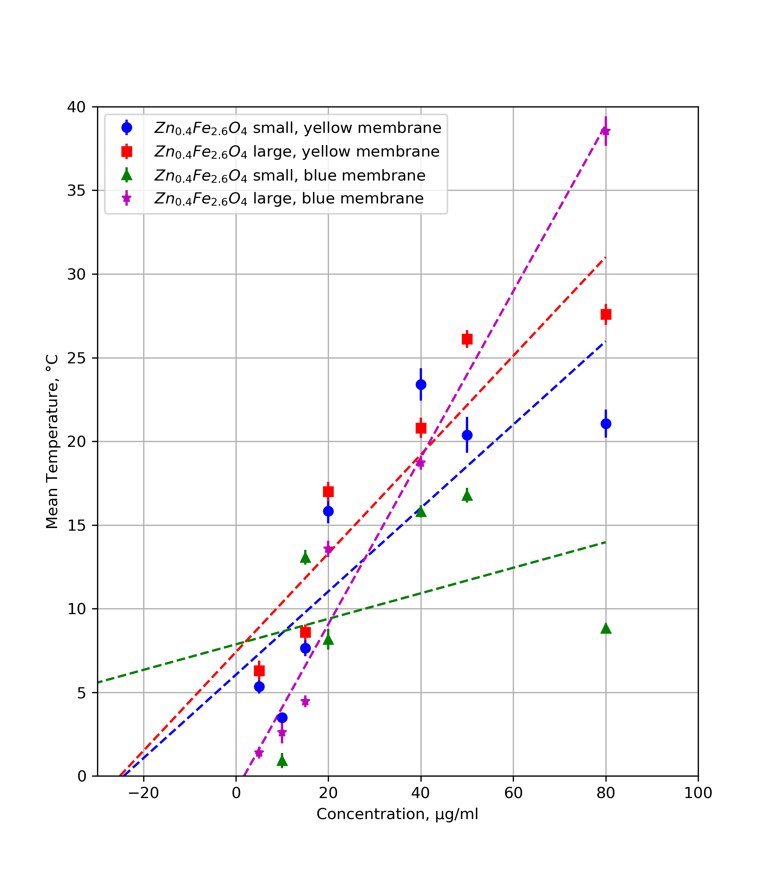


Fig X: *A graph of mean temperature reached by varying sizes of Zinc nanoparticles on different coloured membranes against the concentration of nanoparticles present. Please note the lines of best fit are shown as extrapolated to the left past the data points to illustrate the limit of detection – the concentration at which the mean temperature tends to zero.*

Taking the small Zinc Ferrite nanoparticle results, we can see a fairly large difference in gradients between the two colours of membranes, with the blue membrane mean temperatures increasing more rapidly with concentration than the yellow membrane results. The yellow membrane temperatures seem higher for lower concentrations, and lower than blue for higher concentrations, with an equilibrium reached at roughly 40µg/ml. The gradients and therefore changes in mean temperature with increasing time for the large nanoparticles are remarkably close and noticeably similar to the blue membrane large particle results - perhaps suggesting that the small yellow membrane results may be outliers. We can see here that the blue membrane results are constantly around 3℃ higher than the yellow membrane results, suggesting, given that the Iron Oxide results showed both versions of the trend, that this one is likely to be correct and that yellow membranes absorb more energy from the laser. Again, it is significant that the error estimated on both mean temperature and concentration is far too small, as only a few data points fall on the line of best fit within their uncertainties.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |
|  | Small | | Large | |
| Membrane Colour | Yellow | Blue | Yellow | Blue |
| Limit Of Detection / μg/ml | -24 +/-15 | -103+/-153 | -25+/-11 | 2+/-3 |
| Chi Square | 10.6 | 15.2 | 3.9 | 3.5 |
| Reduced Chi Square | 2.7 | 5.1 | 1.3 | 1.2 |
| Chi-Square Probability | 0.03 | 0.002 | 0.3 | 0.3 |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Zinc Ferrite nanoparticles when exposed to a laser relative to their varying concentration.

Our chi-squared probability values are mostly constantly low here, indicating a bad fit. This is especially true for our small nanoparticles deposited onto a blue membrane. All chi-squared values lie above one, indicating that the null hypothesis that the line of best fit does fit the data should be rejected. The small nanoparticle values indicate a bad fit as they are much greater than one, and the large nanoparticle values show that the fit has not fully captured the trend of the data.

5.3.2: A Logarithmic Trend

On closer inspection of the data points and after significant non-physical limit of detection values, a logarithmic fit was attempted. The data was re plotted using the scipy function scipy.optimize.curvefit on Python and was optimised to the function:

\begin{equation}

y = a\ln{bx} + c

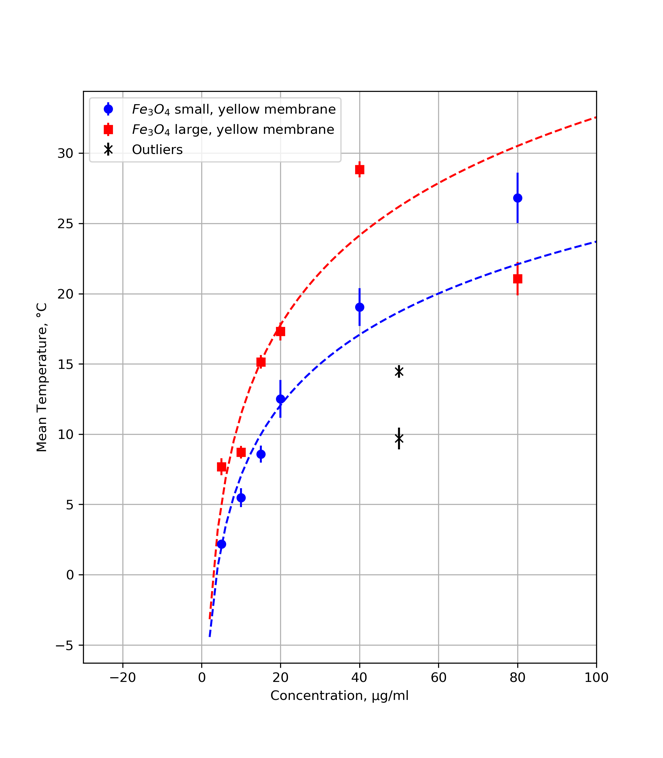
\end{equation}

SUBSUBTIRLE IRON OXIDE

Unfortunately for this fit our blue membrane results had to be excluded as there was too few data points to plot this particular fit for it to be significant. Our data shows the expected positive correlation trend, with the large Iron Oxide nanoparticles heating initially at a greater rate than the smaller ones, and always being located at a greater mean temperature than them. Our limits of detection for this line of best fit are positive and therefore physical, indicating that this is perhaps a better fit than our linear assumption.

Chi squared probability for our large yellow particles is significantly low, indicating a very bad fit to the data points. This is mostly likely due to the variance in the large nanoparticle data. The reduced chi square value is much greater than one, indicating an incorrect fit. It is usual that our chi squared probability is exactly one for our small nanoparticles deposited onto our membrane. This indicates the best fit possible, which is not reflected by the probability, or much by the visual implications of the graph. It is an interesting point here that the data points seem to diverge more at higher concentrations, which indicates that the effect causing our errors is more significant for higher temperatures also.

Outliers are far clearer here, and therefore it was decided that a fit should be attempted with outliers excluded from the data. This fit in shown in Fig. Xb, with the black data points being the removed points.



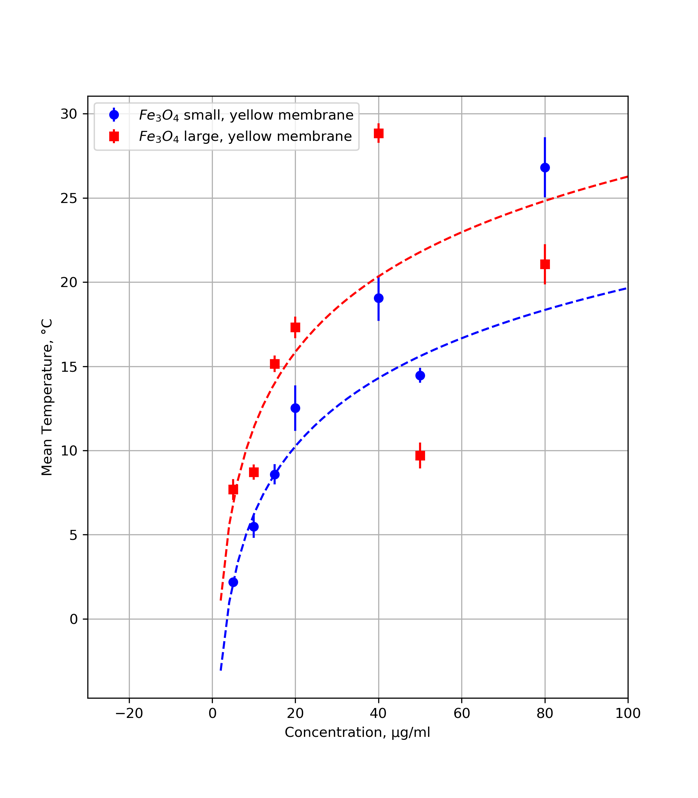


Fig X. just use same as before but words changed and left non outlier removed right outlier removed

The outlier removed results can be seen in Table X. Interestingly our limits of detection are increased, as well as the chi square probability. For our small iron oxide nanoparticles our probability is extremely close to one which could mean that our limit of detection for small iron oxide particles on a yellow membrane is likely close to 3.8 μg/ml. This being said, our reduced chi squared value is close to zero indicating the possibility of overfitting.

|  |  |  |
| --- | --- | --- |
| Nanoparticle size left small right big lol sorry idk how to work these tables | | |
|  |
| Membrane Colour | Yellow | Yellow |
| Limit Of Detection / μg/ml | 3.5 | 1.7 |
| Chi Square | 6.2 | 11.8 |
| Reduced Chi Square | 1.0 | 2.0 |
| Chi-Square Probability | 0.4 | 0.07 |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

|  |  |  |
| --- | --- | --- |
| Nanoparticle size left small right big lol sorry idk how to work these tables | | |
|  |
| Membrane Colour | Yellow | Yellow |
| Limit Of Detection / μg/ml | 3.8 | 2.9 |
| Chi Square | 1.7 | 6.8 |
| Reduced Chi Square | 0.3 | 1.4 |
| Chi-Square Probability | 0.9 | 0.2 |

@sean combine these tables somehow pls this is the one with outliers removed

SUBSUBTIRLE ZINC FERRITE

This fit again gives positive concentration values which indicates that it is perhaps more appropriate. Large and small nanoparticles for each coloured membrane are above and below each other respectively as expected, as large nanoparticles will heat up more per unit concentration than small ones. Blue membrane nanoparticle temperatures seem to lie below yellow membrane values for their respective data pairs, indicating that the blue membranes are absorbing more of the laser wavelength here. We can see a particularly bad fit for the small nanoparticles on the blue membrane which is reflected in its very large reduced chi square value.

We have an extremely small value of chi squared probability for our small yellow data, along with a very high reduced chi square value. The reason for this is unclear, as it seems a better fit than some other curves. Large nanoparticle reducedchi squared values seem very good, although probability is low, most likely due to significant data outliers. With this considered, outliers were removed again and the results are shown below in Fig X and Table X

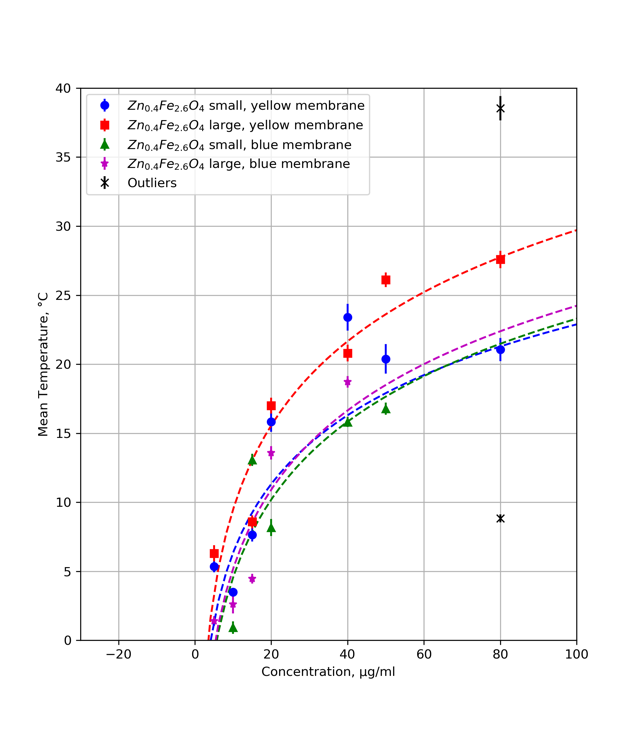
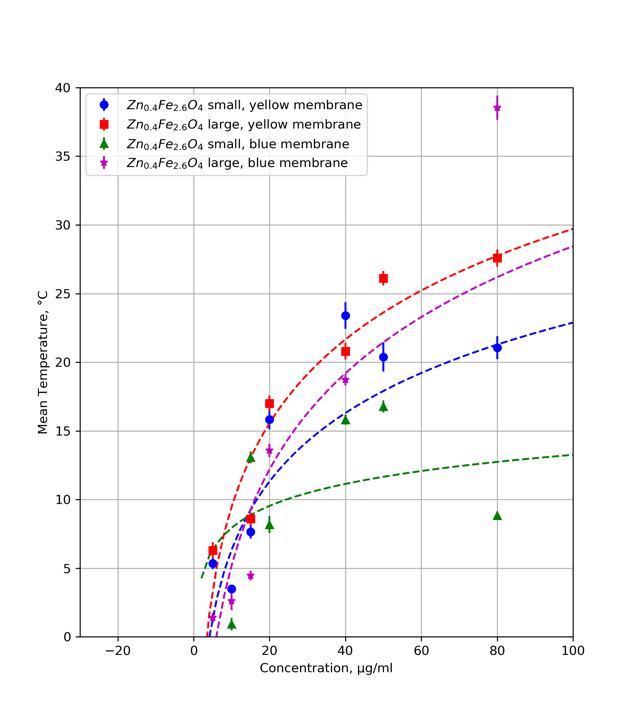


FIG CAPTiONS

USE SYMBOLD FOR RED CHI SQARE AND CHI SQUARE JUST FLOWS NICER

Our blue membrane data produces the only changed fits here. We can see a significant increase in chi squared probability for both, as well as reduced chi squared values tending closer to one. The large blue membrane shows a particularly good result, with chi squared probability equal to one indicating a perfect fit, and reduced chi square value just under one, indicating a fairly good fit with slight underfitting, which makes sense considering the removed data points. In fact, for most chi squared values under one it is most likely due to the fact there are not many data points, and more experimental data needs to be recorded.

Limits of detection are in a good small range for this fit, ranging from 3.4 – 5.7.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |
|  | Small | | Large | |
| Membrane Colour | Yellow | Blue | Yellow | Blue |
| Limit Of Detection / μg/ml | 4.1 | 0.3 | 3.4 | 6.0 |
| Chi Square | 18.5 | 13.8 | 4.5 | 4.2 |
| Reduced Chi Square | 3.1 | 2.8 | 0.9 | 0.8 |
| Chi-Square Probability | 0.005 | 0.02 | 0.5 | 0.5 |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |
|  | Small | | Large | |
| Membrane Colour | Yellow | Blue | Yellow | Blue |
| Limit Of Detection / μg/ml | 4.1 | 5.7 | 3.4 | 5.3 |
| Chi Square | 18.5 | 6.8 | 4.5 | 2.6 |
| Reduced Chi Square | 3.1 | 1.7 | 0.9 | 0.7 |
| Chi-Square Probability | 0.005 | 0.14 | 0.5 | 1.0 |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

5.3.3: A Square Root Trend

Our final fit attempted was a square root trend given by the formula:

y = a\*sqrt(b\*x) + c

Where y is mean temperature data points and x are concentration data points respectively, and a b and c are

SUBSUBTIRLE IRON OXIDE

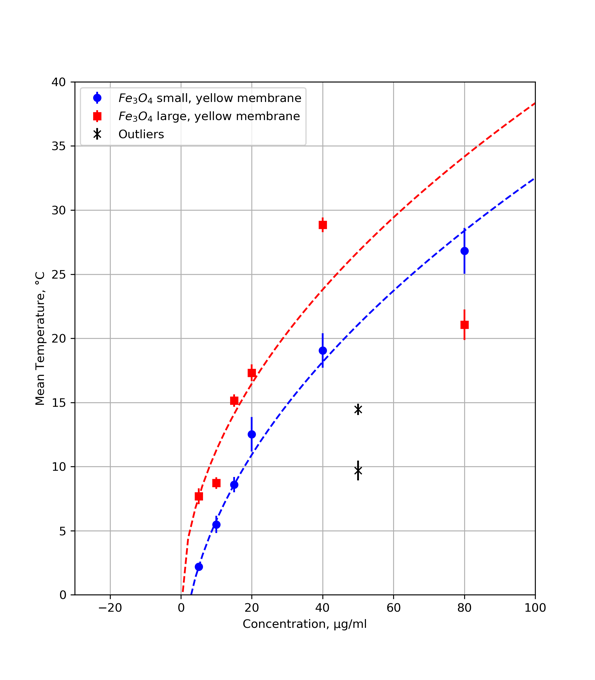
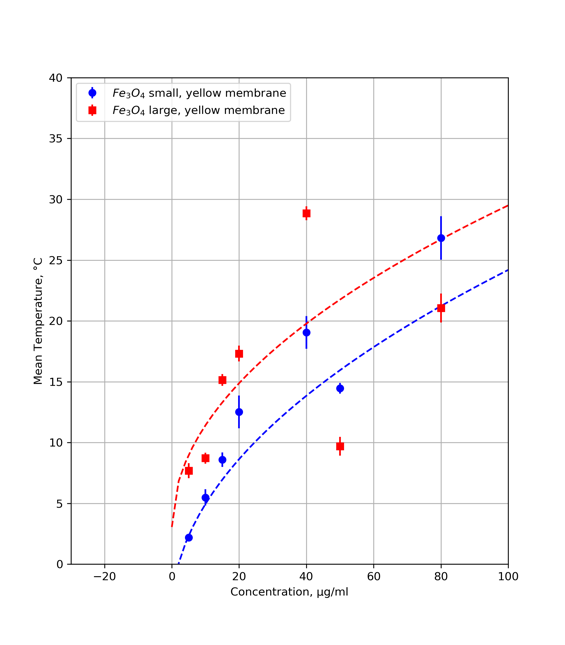


Fig title something or other

Blue membrane data points have again been removed due to there being not enough data to fit the points. We again see the expected positive correlation here, with large nanoparticles mean temperature lying above small for all concentrations as expected. We can see however that we have a negative limit of detection again, indicating a not good fit. This is reflected by an extremely high reduced chi square value and low chi square probability. Unusually, our reduced chi square is exactly one for our small nanoparticles, but is in complete contrast with our very low chi squared probability. Again, this is likely due to outliers, and therefore these were removed again.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |  |
|  | Small | | Large | | |
| Membrane Colour | Yellow | Yellow | |
| Limit Of Detection / μg/ml | 2.0 | -1.3 | |
| Chi Square | 5.7 | 13.5 | |
| Reduced Chi Square | 1.0 | 2.3 | |
| Chi-Square Probability | 0.04 | 0.04 | |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |  |
|  | Small | | Large | | |
| Membrane Colour | Yellow | Yellow | |
| Limit Of Detection / μg/ml | 2.8 | 0.1 | |
| Chi Square | 0.4 | 6.8 | |
| Reduced Chi Square | 0.08 | 1.4 | |
| Chi-Square Probability | 0.08 | 0.09 | |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

Removal of outliers gives visually a much better fit, with mean temperature increasing far more per unit concentration, and both limits of detection now being physical- albeit only just about for the large nanoparticles. Chi square probability is still so low however, that this fit must still be rejected for these data points.

SUBSUBTIRLE ZINC FERRITE

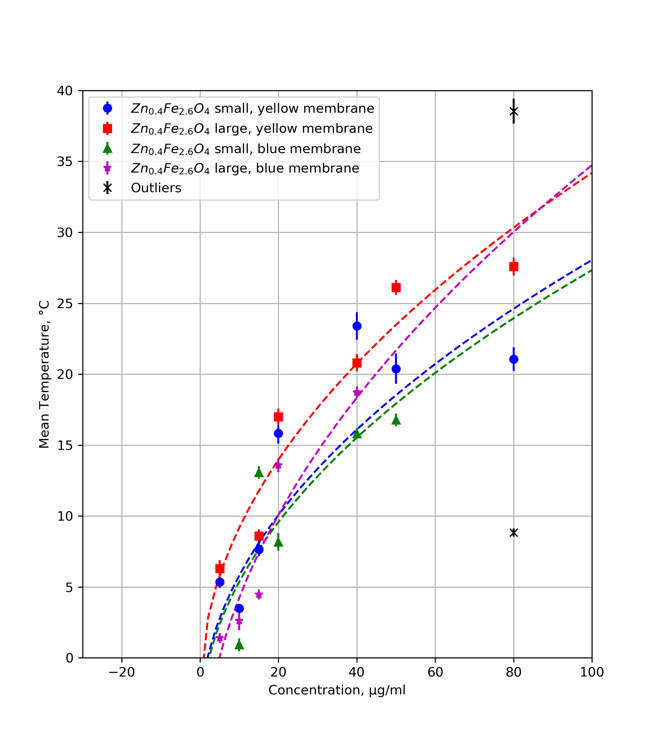
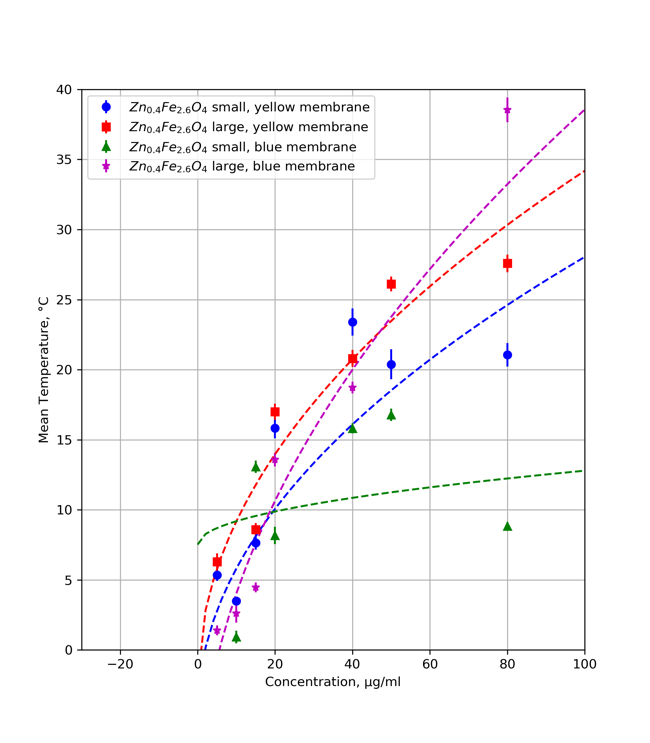


FIG TITLE LOL

We have a clear outlier in the fit here without blue membrane small nanoparticle data. This yields an extremely negative limit of detection value and should be rejected immediately. The other fits show our expected trends, although our large blue membrane data seems to increase unusually rapidly past 40microgram/ml. Probabilities for large nanoparticles are particularly good, although chi square value for blue large is negative so must be rejected. Overall, this fit cannot be said to be conclusive for these data points.

BEFORE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |
|  | Small | | Large | |
| Membrane Colour | Yellow | Blue | Yellow | Blue |
| Limit Of Detection / μg/ml | 1.9 | -203.1 | 0.4 | 5.6 |
| Chi Square | 10.6 | 15.0 | 2.1 | -2.8 |
| Reduced Chi Square | 1.8 | 3.0 | 0.4 | -0.6 |
| Chi-Square Probability | 0.1 | 0.01 | 0.8 | 1.0 |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

AFTER

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Nanoparticle Size | | | |
|  | Small | | Large | |
| Membrane Colour | Yellow | Blue | Yellow | Blue |
| Limit Of Detection / μg/ml | 1.9 | 2.2 | 0.4 | 4.9 |
| Chi Square | 10.6 | 7.8 | 2.1 | 19.4 |
| Reduced Chi Square | 1.8 | 2.0 | 0.4 | 4.8 |
| Chi-Square Probability | 0.1 | 0.1 | 0.8 | 0.0007 |

Table X. Statistical significance of a least squares linear fit fitted to experimental data of change in mean temperature of Iron Oxide nanoparticles when exposed to a laser relative to their varying concentration.

Outlier removal was performed again, with no real significant improvement in chi square values or probabilities, although visually the fit is improved.

ADDITION TO CONCLUSIONS FROM THIS DATA

Chi square values often large indicating overfitting, fitting with noise. Although removal of outliers seemed to visually improve the data, the chi squared probabilities are not often significant enough for them to be considered to tell us that they give good fits. Overall, unfortunately our data is mostly inconclusive due to large noise and not enough data points to draw conclusions.

Chi square values for the logarithmic fit with removal of outliers were most optimal, and so the suggested limits of concentration are suggested as the mean of these values, with the range being the error for each nanoparticle:

2.9 – 3.8 +/- 0.9

3.4 – 5.7 +/- 1.2

SEAN MAYBE DON”T INCLUDE THIS IT IS JUST SO INVALID PROBABLY IDK MESSAGE ME ABOUT IT WHEN YOU GET TO IT

In further runs of the data, more statistically rigorous analysis of outliers must be done, as these were simply guessed by visual analysis of the graph. RIDGE REGRESSION STUFF