Body Mass Index

SL Mathematical Modeling

6/3/2009

Rammy Abssi

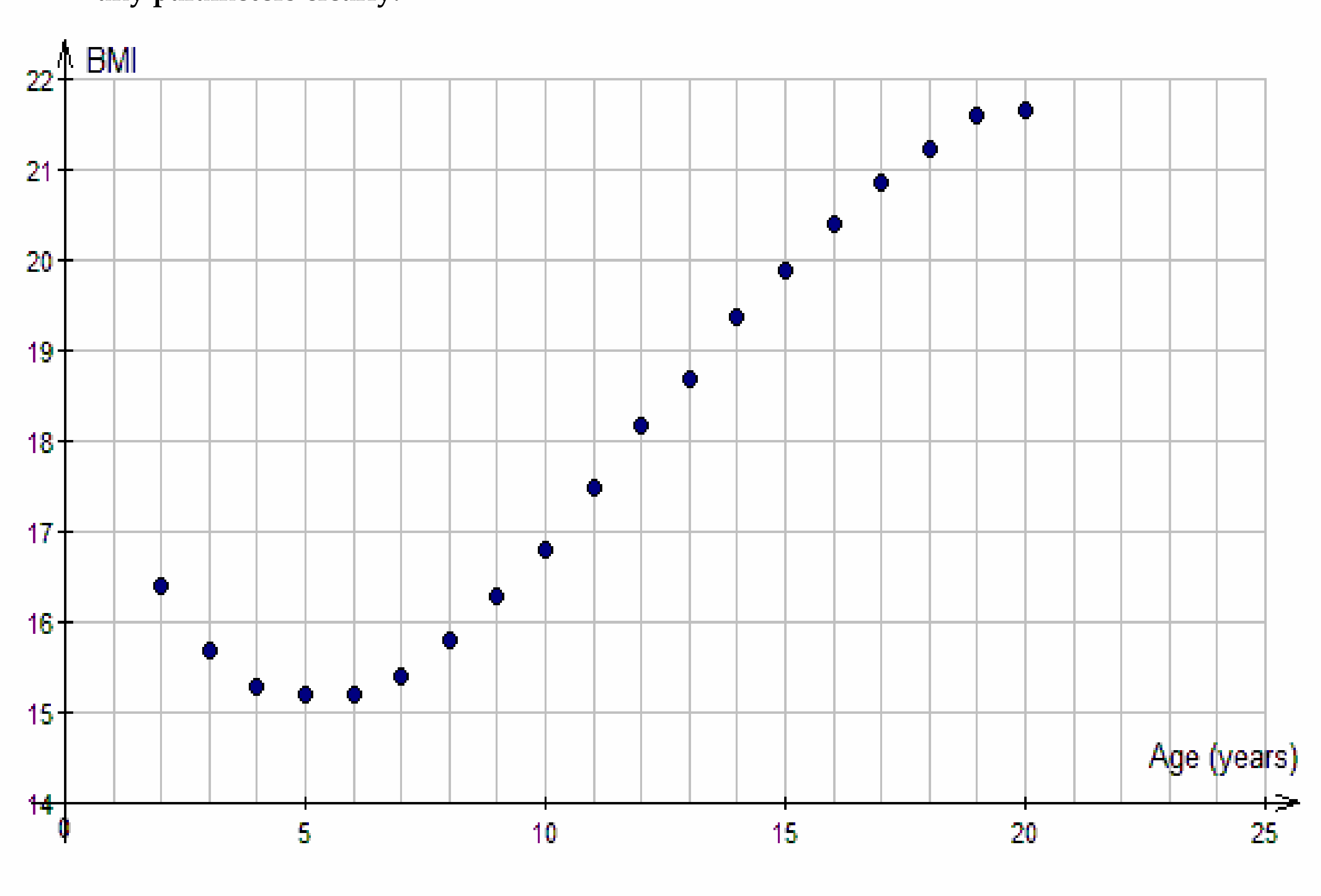
Body mass index (BMI) is a measure of one's body fat. It is calculated by taking one's weight (kg) and dividing by the square of one's height (m).

The table below gives the median BMI for females of different ages in the US in the year 2000.

Age (years)	BMI
2	16.40
3	15.70
4	15.30
5	15.20
6	15.21
7	15.40
8	15.80
9	16.30
10	16.80
11	17.50
12	18.18
13	18.70
14	19.36
15	19.88
16	20.40
17	20.85
18	21.22
19	21.60
20	21.65

(Source: http://www.cdc.gov)

Using technology, plot the data points on a graph. Define all variables used and state any parameters clearly.



In the graph above, the variables shown are age in years and BMI. In this case, the x axis is the age of the female in years and the y axis is the body mass index of the females.

The parameters are also known as the domain and range. The parameters of the plotted data are shown below:

$D: 2 \leq x \leq 20$

$R: 15.2 \leq y \leq 21.65$

The domain and range can also be seen clearly from the table. The x values range from 2 to 20, and the y values range from 15.2 to 21.65.

What type of function models the behavior of the graph? Explain why you chose this function. Create an equation (a model) that fits the graph.

Throughout this study, many different types of function models were experimented with. However, the one which models the behavior of the graph the best would be a sine function.

A simple sine function looks like:

$$y = \sin(x)$$

Then this sine function would need to be manipulated to fit the data. This function was chosen because there is an obvious trend in the data that shows a curve. The sine function can easily be manipulated to fit the curve. The following equation shows how to manipulate the sine function:

$$y = a\sin(bx + c) + d$$

In this case all of the variables, which include *a, b, c,* and *d* would need to be manipulated to fit the data.

First, we must understand what each variable does:

Variable	Affect if Variable is Changed
a	Affects the amplitude of the sine function
b	Affects the period of the sine function
c	A horizontal translation of the sine function
d	A vertical translation of the sine function

Looking at the plotted data intimately and trying many different changes on the variables I came to the conclusion of:

a = 3

b = .2

c = -2.8

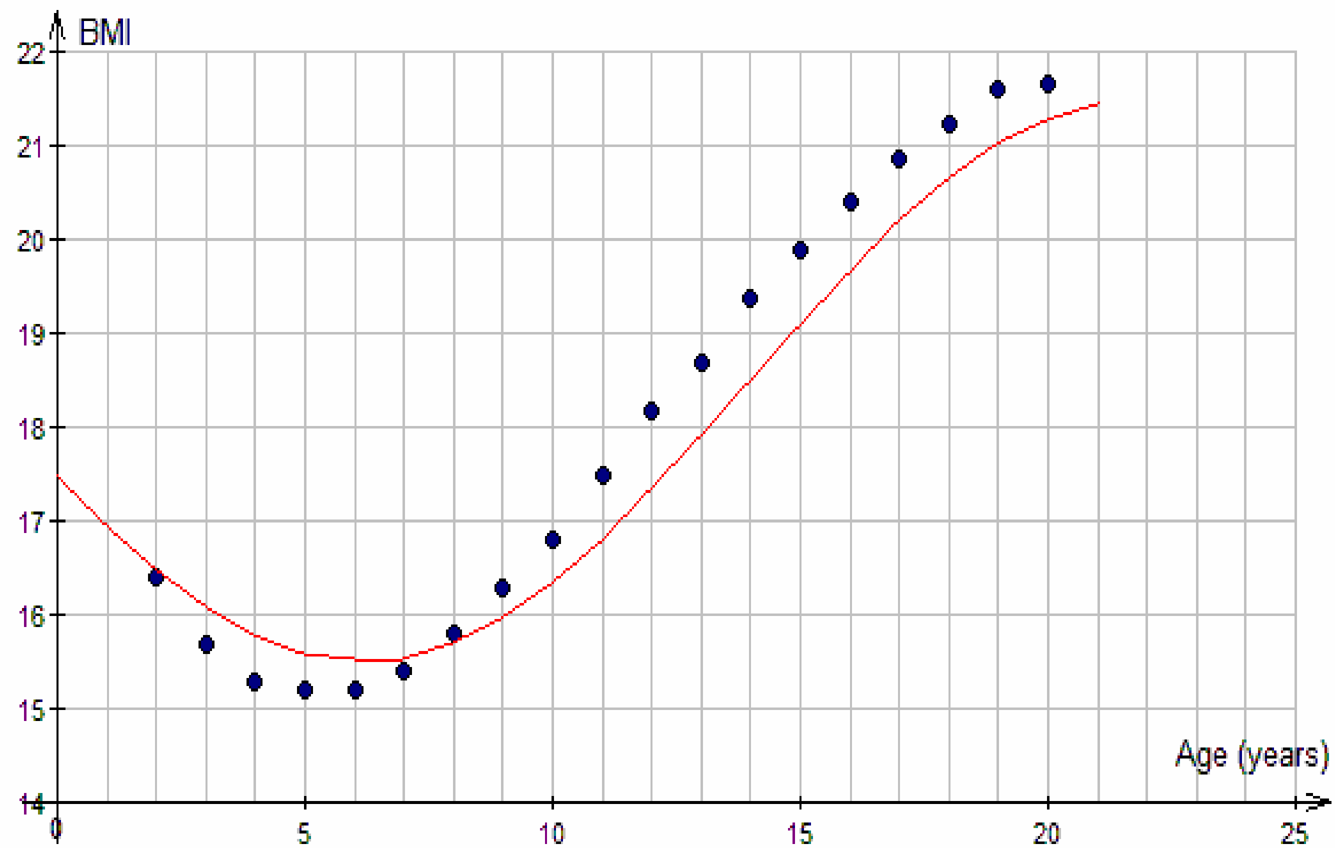
d = 18.5

Therefore, the equation would look like:

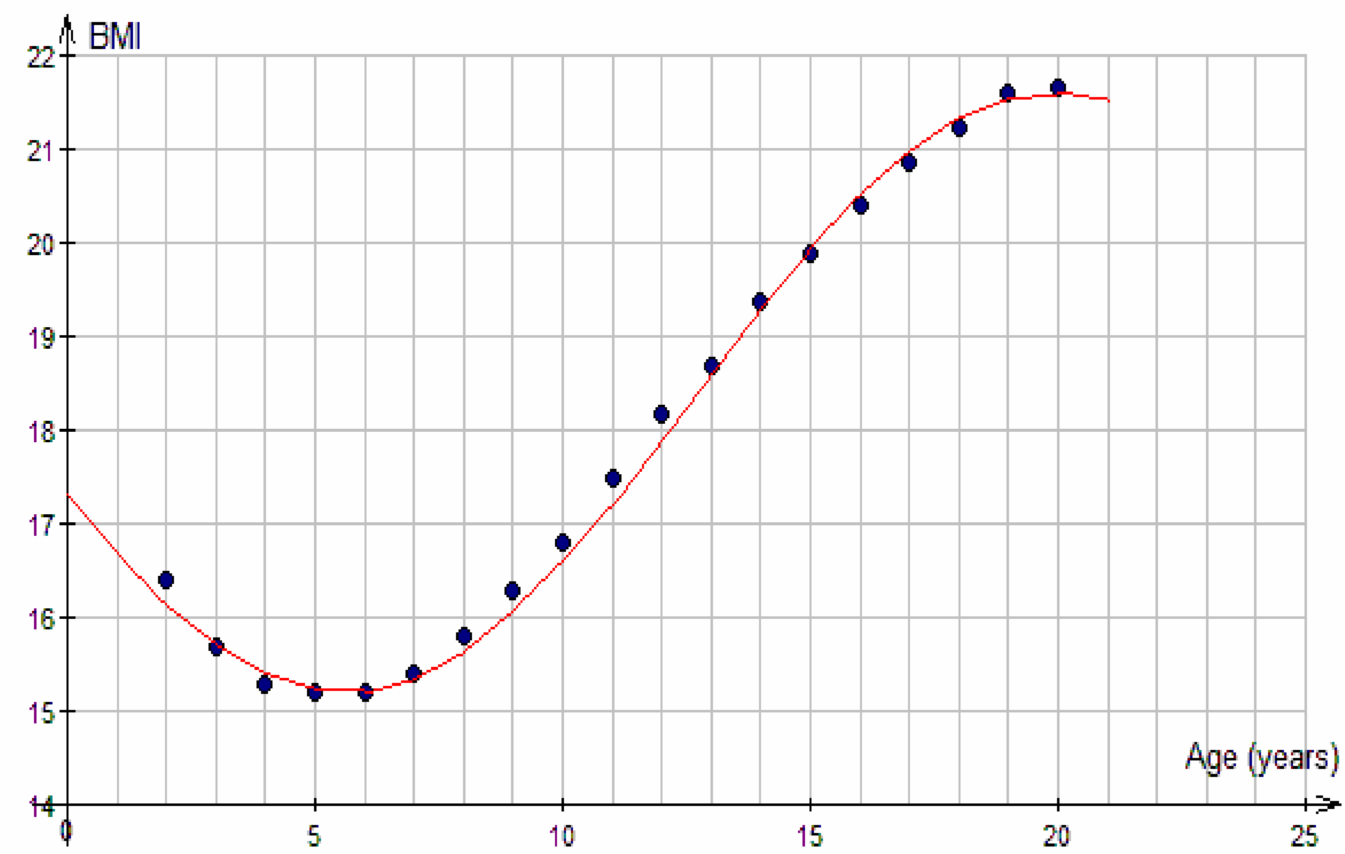
$y = 3\sin(0.2x - 2.8) + 18.5$

The equation above was graphed along with the plotted data and there was similarity between the two.

On a new set of axes, draw your model function and the original graph. Comment on any differences. Refine your model if necessary.



The blue dotted points are the data points from the table on the first page. The red line is the sine function arrived at earlier. Obviously, the line is not perfectly in sync with the plotted data; however, it is reasonably close. After graphing both on the same axes some clear differences can be seen more clearly. For example, first the minimum of the function is higher than the minimum of the data. This can be improved by changing the variable *d*. By changing *d* slightly the whole sine function can move down however many units you increase *d* by. Also the sine function curves off of the plotted data. This can be improved by increasing the frequency of the sine function very slightly. However, when changing the period, the other variables would need to be manipulated slightly to accompany the change.



The graph on the next page is the refined graph. Slight changes were made to make the function more accurate to the data.

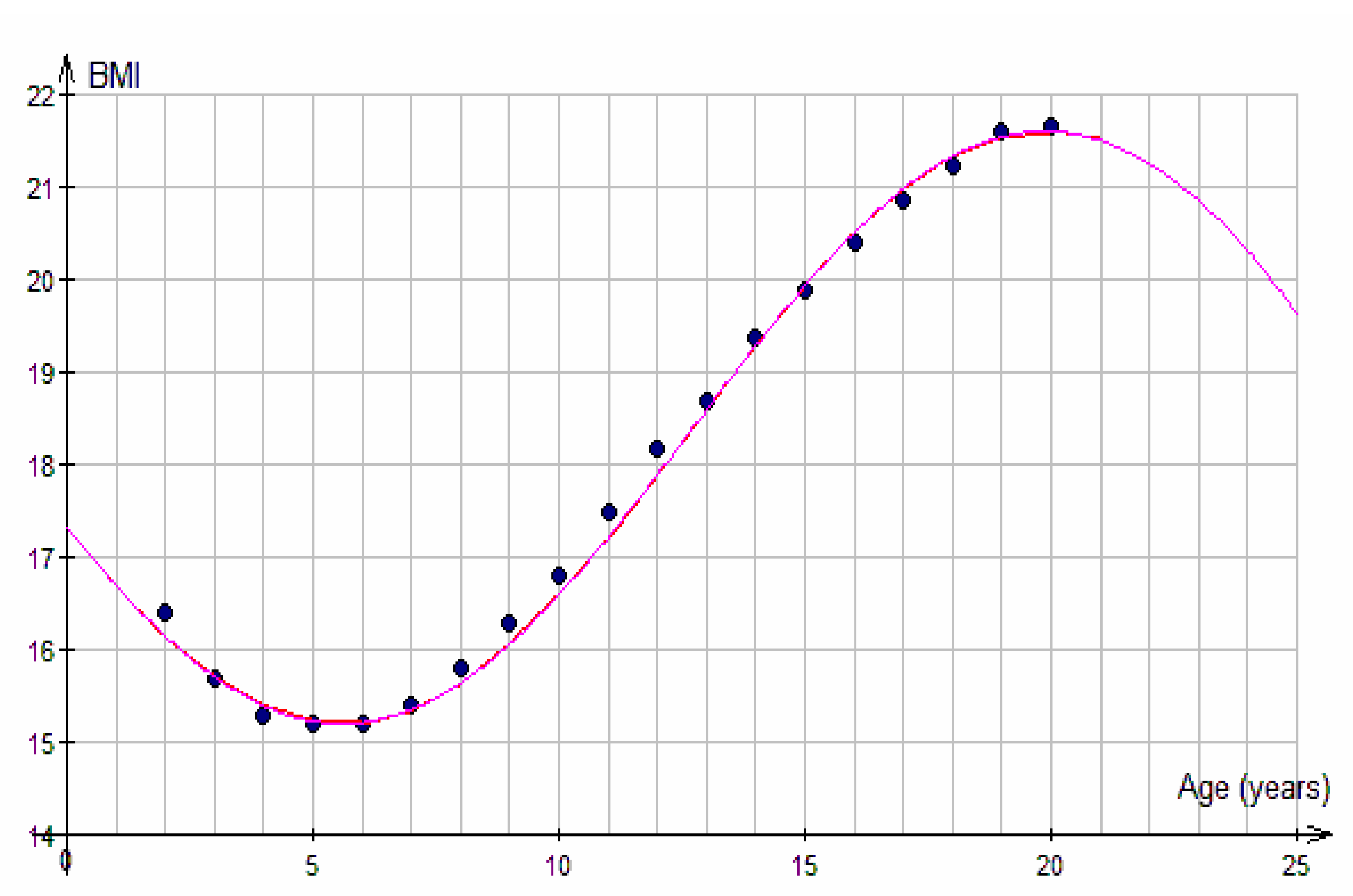
The graph above shows the plotted data points with the refined and more accurate model. The function of this model was:

$y = 3.1\sin(0.22x - 2.8) + 18.4$

As you can see here, this function is much more accurate. The line of the function is much more in sync with the data points that are plotted.

Use technology to find another function that models the data. On a new set of axes, draw your model function and the function you found using technology. Comment on any differences.

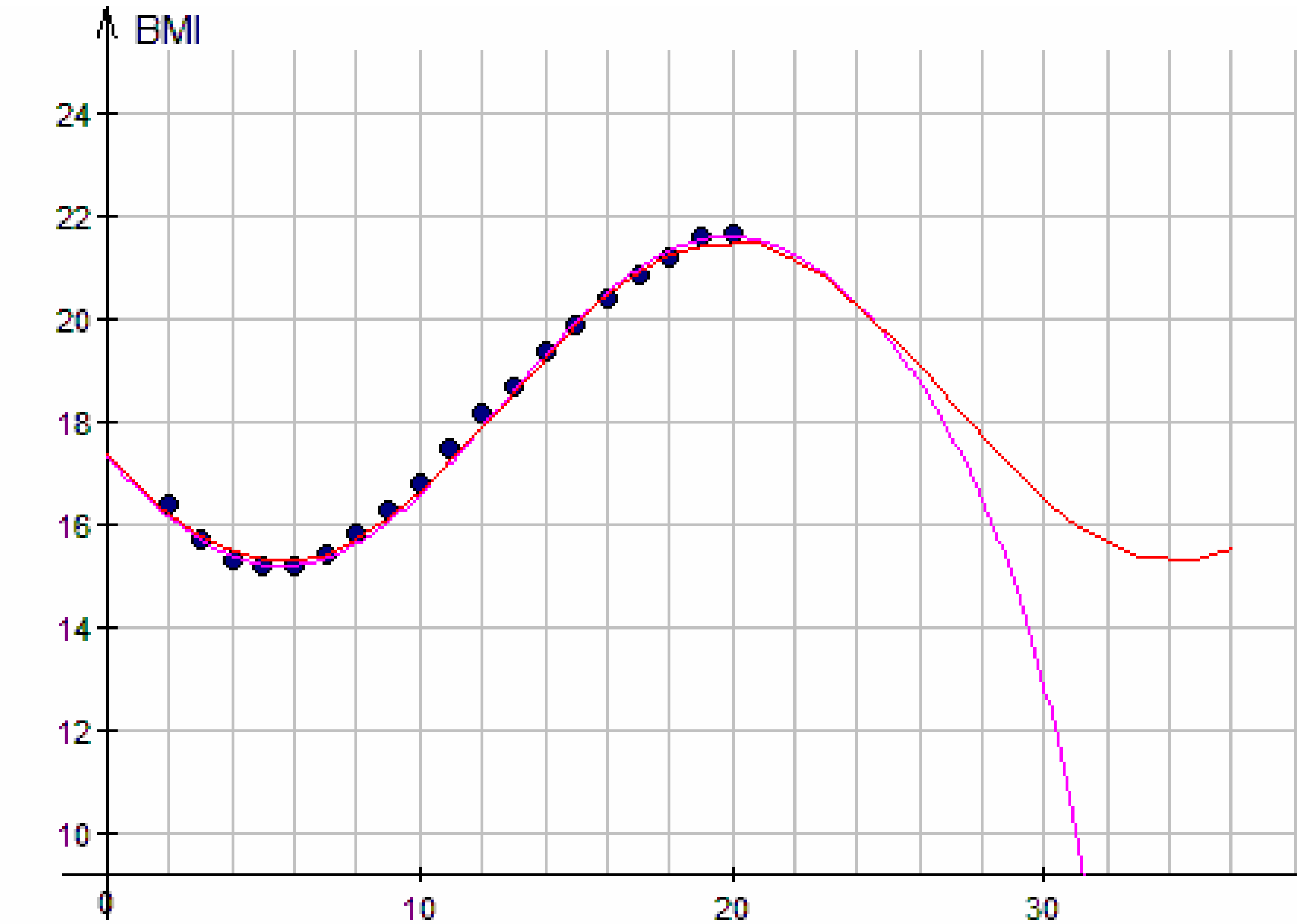
Even though the function arrived at above was quite accurate the use of technology can develop a function which is also accurate. To do so, a program by the name of ZGrapher is used. First, the points from the raw data are plotted. After this there is a tab at the top of the program called 'Calculus.' Under this tab you can use the regression tool to find the best fit cure.



The graph above shows three different factors. The blue points are the data points of the BMI table. The red line shows the earlier model which was arrived at, and the pink line shows the best fit curve from the program. As you can see, the red line and pink line are almost identical; therefore, the model arrived at earlier would be very accurate. However, one clear and very important difference would be how the function changes as the scale increases. The two functions act very differently after 20 years on the x axis.

Use your model to estimate the BMI of a 30-year-old woman in the US. Discuss the reasonableness of your answer.

As arrived at earlier, both functions act very differently after about 20 years. Therefore, both functions will be graphed on a different axis to estimate the BMI of a 30-year-old woman. To do so, the x axis would need to be expanded and when the line intersects with



x = 30 the BMI of that woman can be estimated.

The graph above shows very clearly how the BMI for women was estimated at the age of 30. The red line is my model function, where as the pink line was the regression function recognized from the technology.

For my model, the estimated BMI at the age of 30 is about **16.5**.

For the regression model, the estimated BMI at the age of 30 is about **13**.

The results for both of the functions are unreasonable. Even though there is a possibility where this could be true, one has to interpret what this means.

The BMI is calculated by taking one's weight in kilograms and dividing it by the square of one's height in meters. After about the age of 20, the growth in height of women generally stops. If this is the case, the denominator in the calculation would generally stay the same. Therefore, the BMI would depend on the weight of the woman. If a woman's weight increases than the BMI would increase.

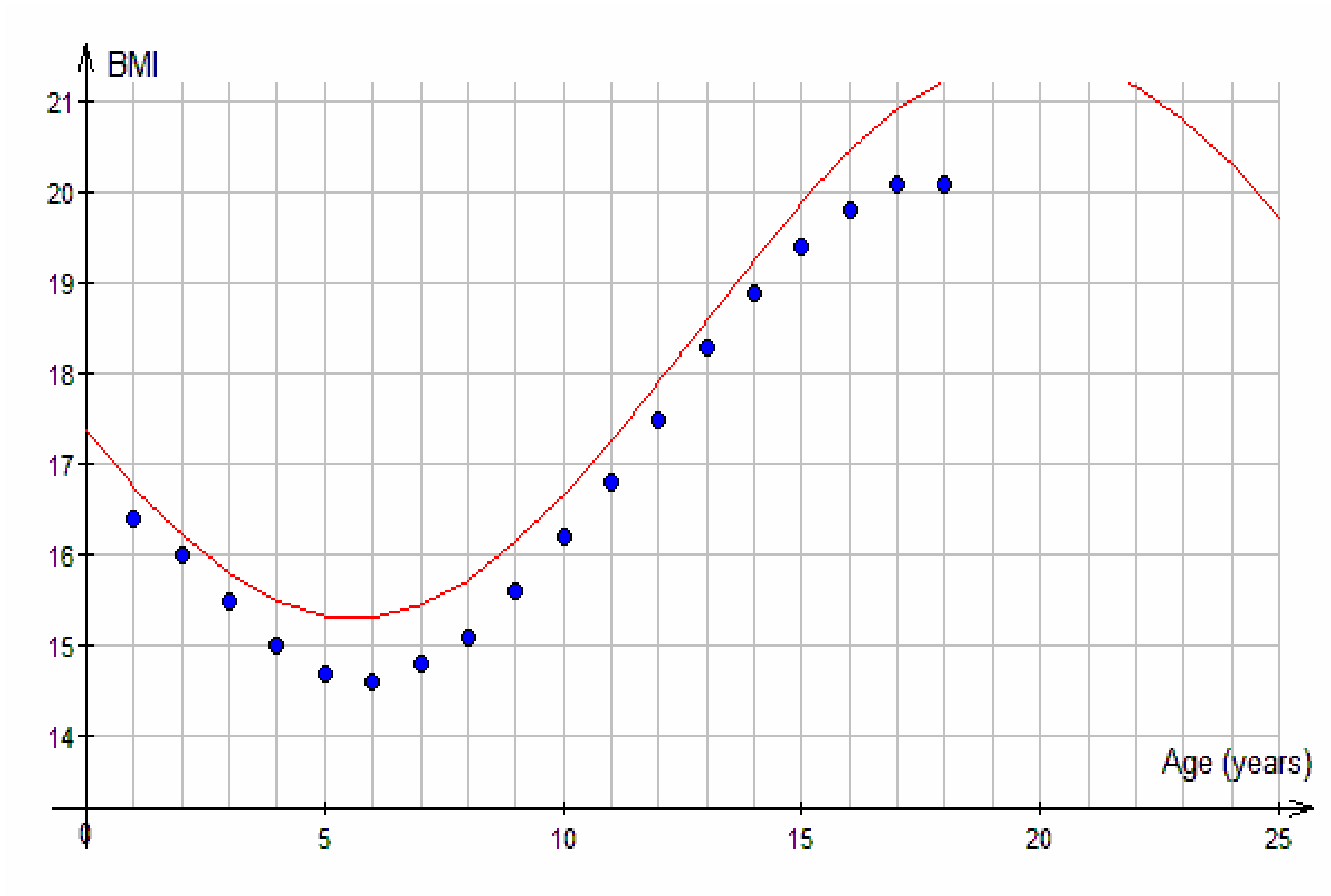
Relating this back to the functions and estimated values for the BMI of a 30-year-old woman, it is seen that the BMI decreases dramatically. This would indicate that the woman went under a massive weight loss after the age of twenty. This is exactly why the estimated values for the BMI of a woman at the age of 30 is unreasonable.

Use the internet to find BMI data for females from another country. Does your model also fit this data? If not, what changes would you need to make? Discuss any limitations to your model.

The table below gives the average BMI for females of different ages in Shaanxi, China.

Age (years)	BMI
1	16.4
2	16.0
3	15.5
4	15.0
5	14.7
6	14.6
7	14.8
8	15.1
9	15.6
10	16.2
11	16.8
12	17.5
13	18.3
14	18.9
15	19.4
16	19.8
17	20.1
18	20.2

(Source: www.ijbs.org/User/ContentFullText.aspx?VolumeNO=1&StartPage=57&Type=pdf)



The red line in this graph represents the same model function as before, and the blue data points are the BMI data points for girls ages 0 to 18 in China. As you can see the data follows a very similar trend to the model; however, the overall BMI of women in China is less than the overall BMI of women in the US. To make the model fit this data the whole sine function model would need to be moved down about half of a unit.

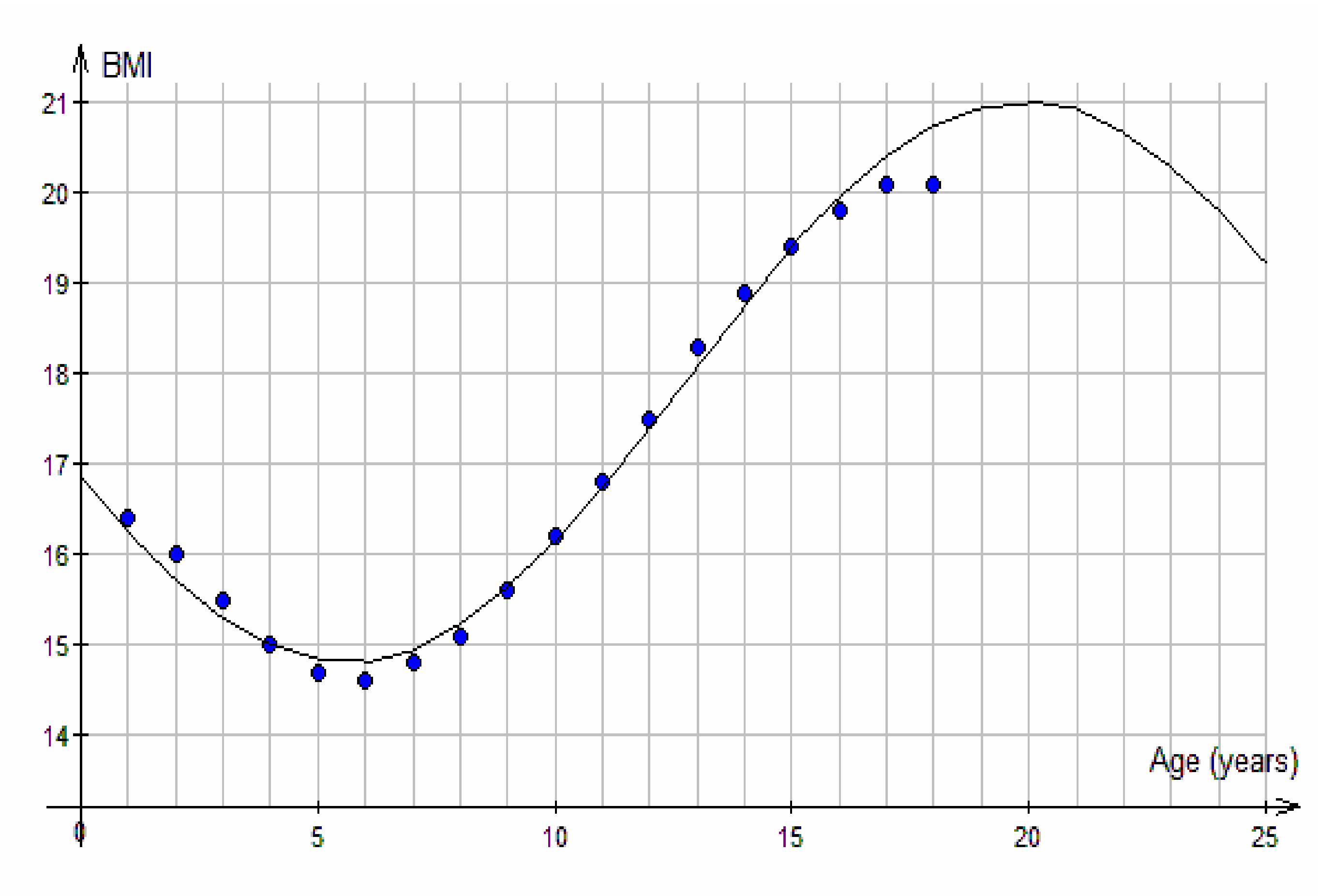
The model function which was concluded earlier looked like the following:

$y = 3.1\sin(0.22x - 2.8) + 18.4$

After, refining this function to fit the data from china the variable *d* was decreased by half of a unit:

$y = 3.1\sin(0.22x - 2.8) + 17.9$

By decreasing the variable *d* by half of a unit, the whole model function undergoes a



vertical translation of 0.5 units downwards.

As you can see in the graph above, the refined function fits the data on women from China much better.

Overall, the models used in this investigation were closely related to the data; however, there is a very limiting problem with the sine model used. The main limitation was seen when estimating the BMI of a 30-year-old woman. Due to the fact that the sine function is recurring after the maximum point (which was around the age of 20), the y value begins to decrease. This is a very important limitation because if the BMI of a woman needs to be estimated the information would neither be reasonable nor reliable.