

THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics

Mathematical Modeling Project Team
mathmodel@math.cuhk.edu.hk

Modeling with R Shiny (General Fitting for XY data)

Last updated: March 11, 2026

1 Introduction to the General Fitting (XY) R Shiny tool

一般擬合(XY) R Shiny 工具簡介

R is a powerful statistical programming language that allows us to perform data analysis, which is useful for mathematical modeling. Even better, we have set up a set of online tools using R Shiny so that you can perform various mathematical modelling tasks without any coding background.

R 是一種功能強大的統計編程語言，讓我們能夠進行數據分析，這對數學建模非常有用。更理想的是，我們利用R Shiny 建立了一系列網上工具，讓大家即使沒有任何編程背景，也能夠執行各種數學建模任務。

In particular, a common task in mathematical modelling is to perform nonlinear regression based on some data points $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Other than our Nonlinear Regression (XY) tool, we have the General Fitting (XY) tool for you to input your own functions for data fitting. Here is the guideline on how to use our General Fitting (XY) R Shiny tool.

在數學建模中，一項常見任務是根據若干數據點 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 進行非線性迴歸。除了我們的非線性迴歸(XY) 工具外，我們還提供一般擬合(XY) 工具，讓使用者可以自行輸入函數進行數據擬合。以下是有關如何使用我們的一般擬合(XY) R Shiny 工具的指引。

2 Using the General Fitting (XY) R Shiny tool

使用一般擬合(XY) R Shiny 工具

To use the tool, you can go to our main website:
要使用此工具，可前往我們的主網站：

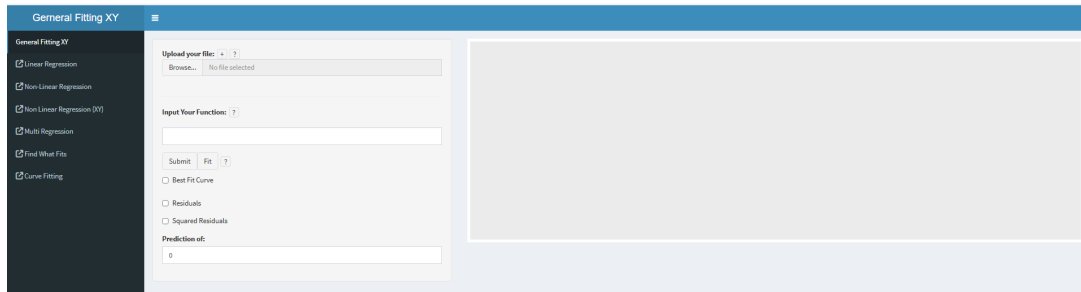
<https://www.math.cuhk.edu.hk/app/mathmodel/tool.html>

for the list of tools, or simply go to
瀏覽工具列表，或直接前往

<https://mathmodelcuhk.shinyapps.io/general-fitting/>

for the General Fitting (XY) R Shiny tool.
使用一般擬合(XY) R Shiny 工具。

After getting into the website, you can see the following:
進入網站後，你會看到以下畫面：



Don't worry if you think this is complicated! We will guide you step by step, and you can do nonlinear regression using our R Shiny tool.
如果你覺得這個界面有點複雜，不用擔心！我們會逐步引導你，讓你可以使用我們的R Shiny 工具進行非線性迴歸。

2.1 Step 1: Data input

步驟一：數據輸入

At the top left-hand corner, you can see “Upload your file”, this is where we input the data. There are three different buttons.

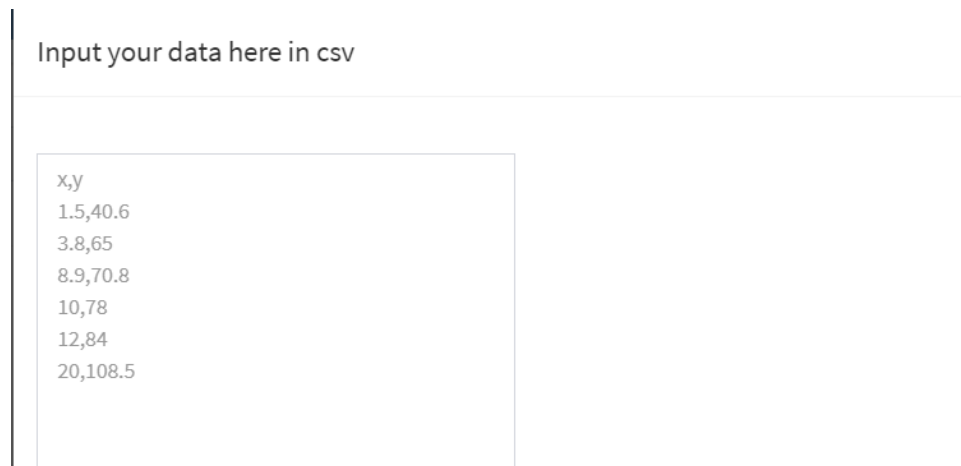
在左上角，你會看到「Upload your file」，這就是我們輸入數據的地方。這裡有三個不同的按鈕。

Firstly, the “?” gives you the general guide on data input.

首先，「？」按鈕會提供關於數據輸入的一般指南。

Secondly, the “+” allows you to input data directly. After clicking it, you can see the following:

其次，「+」按鈕讓你可以直接輸入數據。點擊後，你會看到以下畫面：



The screenshot shows a web interface for data input. At the top, it says "Input your data here in csv". Below this is a text area containing the following CSV data:

```
x,y
1.5,40.6
3.8,65
8.9,70.8
10,78
12,84
20,108.5
```

On the first line, you can input the names of the variables for the x-axis and y-axis respectively. For example, we can define the x-axis as “Month” and the y-axis as “Average hours of daylight” to see the correlation between them. Therefore, on the first line we type “Month, Average hours of daylight”

在第一行，你可以分別輸入x軸和y軸的變量名稱。例如，我們可以將x軸定義為「月份」，y軸定義為「平均日照時數」，以觀察兩者之間的相關性。因此，在第一行我們輸入「Month, Average hours of daylight」。

For the following lines, you can simply type in the data points respectively. You should be able to have a result that looks like this:

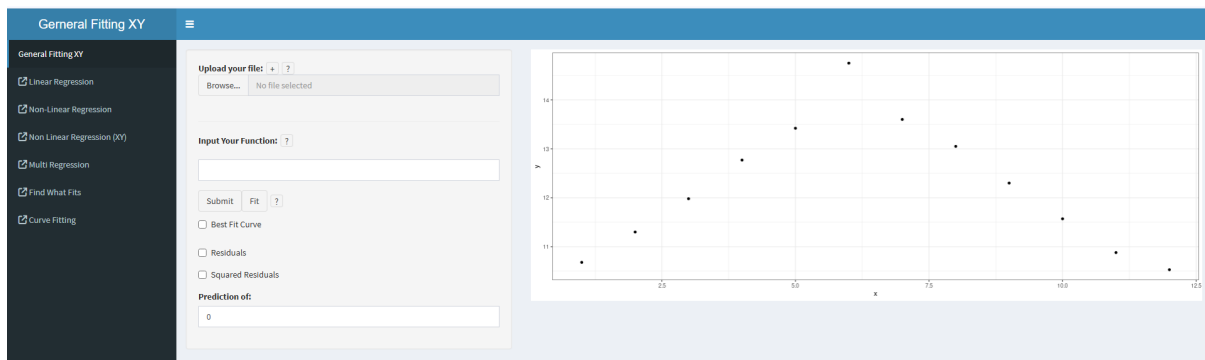
在接下來的行數，你可以分別輸入各個數據點。你應該會得到類似這樣的結果：

Input your data here in csv

```
Month, Average hours of daylight
1,10.68
2,11.3
3,11.98
4,12.77
5,13.42
6,14.75
7,13.6
8,13.05
9,12.3
10,11.57
11,10.88
12,10.53
```

Scroll down and press “submit”. Your data points should be shown on the screen like this:

向下滾動並按下「submit」，你的數據點應該會在屏幕上顯示，如下所示：



You can also upload an Excel file to the website directly. Firstly, you will need to prepare your Excel file like this:

你也可以直接將Excel 檔案上傳到網站。首先，你需要準備好Excel 檔案，格式如下：

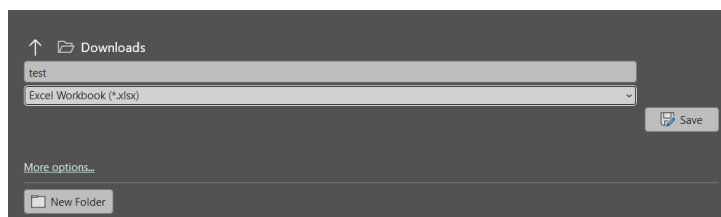
	A	B
1	Time	Hours of Daylight (hrs)
2	1	10.68
3	2	11.3
4	3	11.98
5	4	12.77
6	5	13.42
7	6	14.75
8	7	13.6
9	8	13.05
10	9	12.3
11	10	11.57
12	11	10.88
13	12	10.53
14		

Then go back to our website, press the “Browse...” button and look for the file stored on the computer. After uploading, you will see your data points shown on the graph.

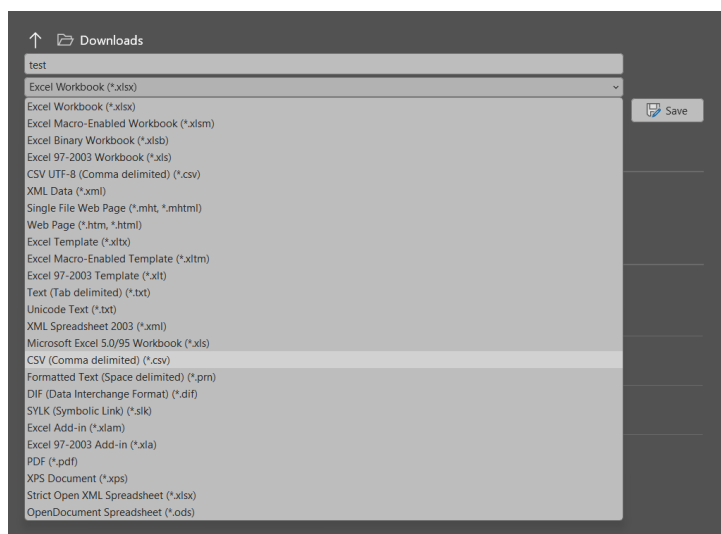
然後返回我們的網站，按下「Browse...」按鈕，尋找儲存在電腦上的檔案。上傳完成後，你會看到數據點顯示在圖表上。

If you face any problem regarding the upload, you may need to save your Excel file in .csv format. To do so, you can go to file → save as, and you can look for the following:

如果在上傳時遇到任何問題，你可能需要將Excel 檔案儲存為.csv 格式。方法是前往「檔案」→「另存新檔」，然後尋找以下選項：



Expand the dropdown list, and look for the .csv format.
展開下拉列表，尋找.csv 格式。

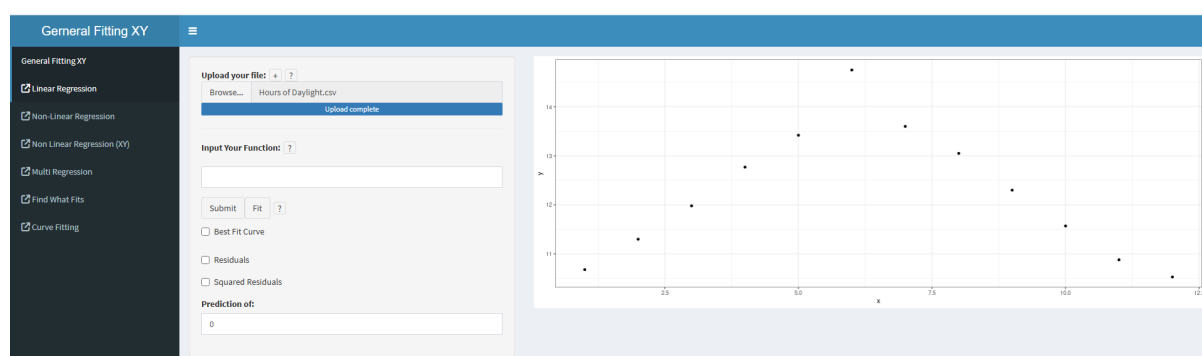


After that, you can upload the file onto our website, and it will work normally.
完成後，你可以將檔案上傳到我們的網站，它就能正常運作了。

2.2 Step 2: Data analysis 步驟二：數據分析

After you have inputted the dataset, now we can perform linear or nonlinear regression easily. Try to click different buttons on the page and now we will briefly explain their usage.

輸入數據集後，我們現在可以輕鬆地進行線性或非線性迴歸。嘗試點擊頁面上的不同按鈕，現在我們將簡要說明它們的用途。



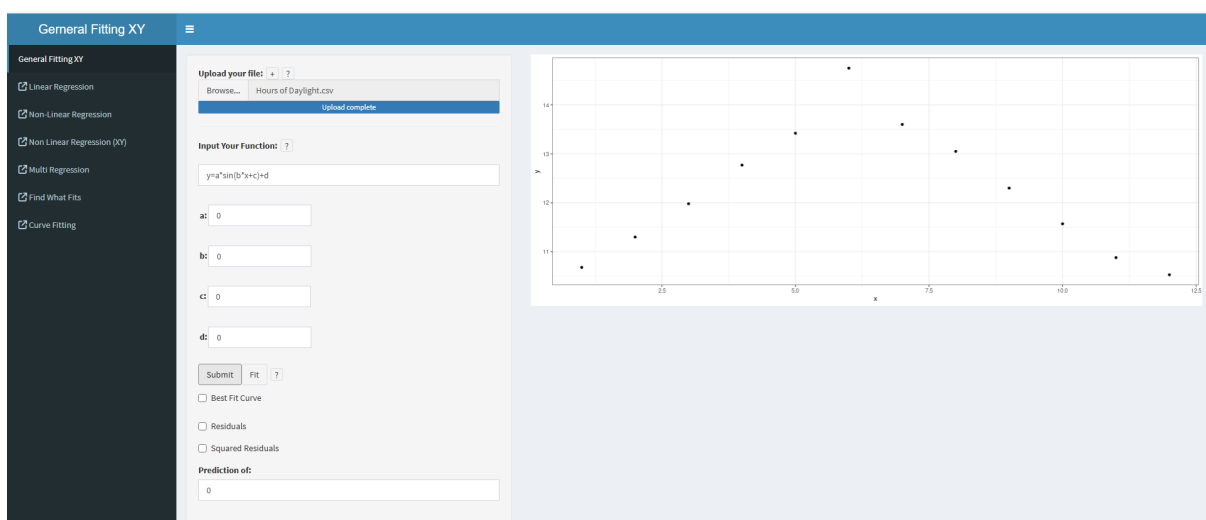
If you have any questions about the tool, you can always try to click different boxes with “?” and there will be a short introduction.

如果你對這個工具有任何疑問，可以隨時點擊帶有「？」的方框，那裡會有簡短的介绍。

2.2.1 Input Your Function 輸入你的函數

In the box below, you can input the function that you want to use for the data fitting. In our case, since the hours of daylight are periodic throughout seasons, we try to model it with a sinusoidal function. Therefore, we will type $y = a * \sin(b * x + c) + d$ into the box. Please note that here a, b, c, d are parameters that are to be determined under fitting. Then we will click “Submit” to enter the function. And you will see there are four boxes to be filled.

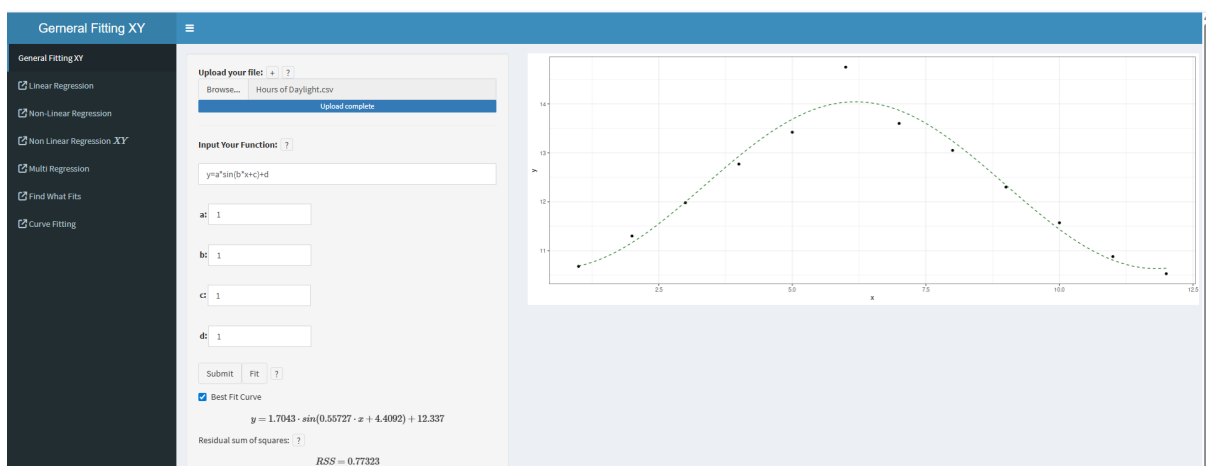
在下方的方框中，你可以輸入想要用於數據擬合的函數。以我們的情況為例，由於日照時數隨季節變化而呈現週期性，我們嘗試用正弦函數來建模。因此，我們會在方框中輸入 $y = a * \sin(b * x + c) + d$ 。請注意，這裡的 a, b, c, d 是在擬合過程中需要確定的參數。然後我們點擊「Submit」輸入函數，你會看到出現了四個需要填寫的方框。



The screenshot shows the 'General Fitting XY' interface. On the left, a sidebar lists various fitting methods, with 'Curve Fitting' selected. The main area is divided into two panels. The left panel, titled 'Input Your Function', contains a text input field with the formula $y = a * \sin(b * x + c) + d$. Below this are four input fields for parameters a , b , c , and d , each currently set to 0. There are 'Submit' and 'Fit' buttons, and a 'Best Fit Curve' checkbox which is currently unchecked. The right panel shows a scatter plot of data points on a grid, representing the 'Hours of Daylight.csv' dataset.

This is for the initial guess for the function (to know more, you can look for the gradient descent method online, but let's ignore it for now). We can simply type in 1, 1, 1, 1 for all values. After that, we click “Fit” and then check the “Best Fit Curve”. You will see the following:

這是為函數提供初始估計值（想了解更多，可以在網上搜尋梯度下降法，但我們暫時先忽略這部分）。我們可以簡單地為所有值輸入1, 1, 1, 1。完成後，點擊「Fit」，然後勾選「Best Fit Curve」。你會看到以下結果：



The screenshot shows the 'General Fitting XY' interface after the 'Fit' step. The function $y = a * \sin(b * x + c) + d$ is still entered. The parameters a , b , c , and d are now all set to 1. The 'Fit' button is highlighted, and the 'Best Fit Curve' checkbox is checked. Below the input fields, the fitted equation is displayed: $y = 1.7043 - \sin(0.55727 * x + 4.4092) + 12.337$. The residual sum of squares is shown as $RSS = 0.77323$. The scatter plot on the right now shows a green dashed sinusoidal curve fitted to the data points.

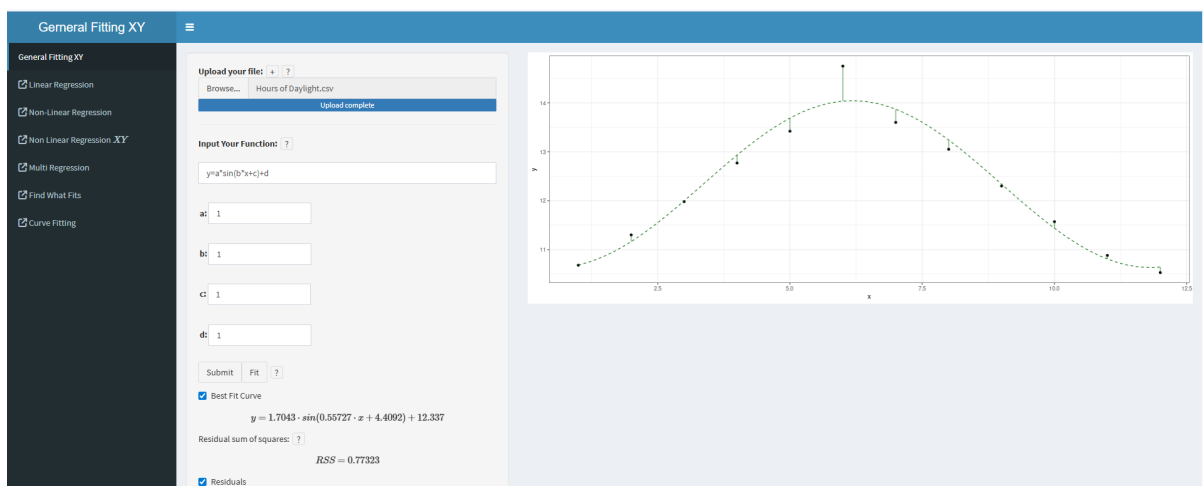
Together with the graph, the tools also tell you different related information, such as the equation of the best fit curve and the residual sum of squares. The residual sum of squares measures how well this curve can approximate the dataset we have. You can use this to compare which model is better for the regression, too. (Please note that a low residual sum of squares does not always imply the model we are using is a good one.)

除了圖表之外，工具還會顯示各種相關資訊，例如最佳擬合曲線的方程和殘差平方和。殘差平方用於衡量這條曲線對我們所擁有數據集的近似程度。你也可以用這個數值來比較哪個模型更適合進行迴歸分析。（請注意，殘差平方和較低並不總是表示我們使用的模型是好的。）

2.2.2 Residuals 殘差

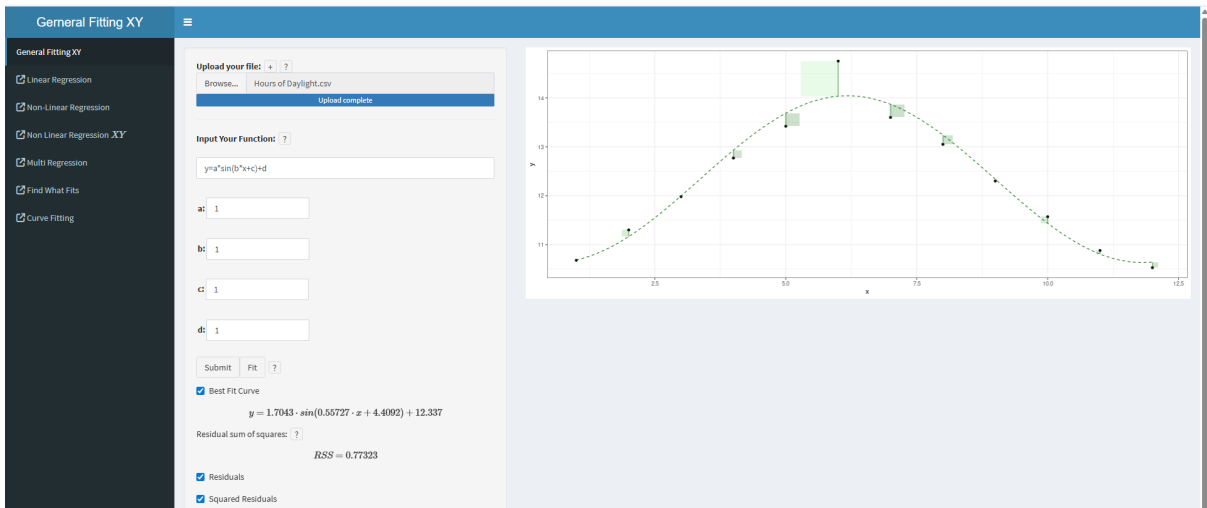
This gives us a set of vertical lines, showing the difference between each of our original data points and the regression curve.

這會顯示一組垂直線，表示每個原始數據點與迴歸曲線之間的差異。



2.2.3 Squared Residuals 平方殘差

This function visualizes the square of residuals as the area of the rectangles. In fact, it is originally a square with side length equal to the residual, just that in the graph, the vertical and horizontal scales are different, and that makes the shape look like a rectangle. 此功能將殘差的平方可視化為矩形的面積。實際上，它原本是一個邊長等於殘差的正方形，只是由於圖表中垂直和水平的比例尺不同，使得形狀看起來像一個矩形。

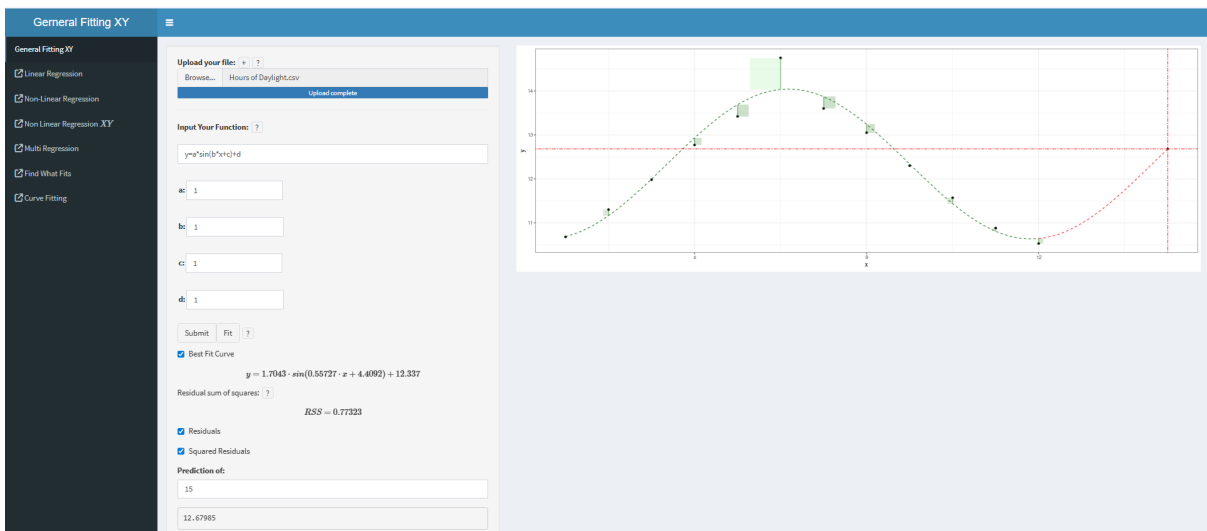


2.2.4 Prediction 預測

By typing any number into the box, we can know the corresponding prediction of the y-coordinate according to our regression curve. You can also see how the regression line is extended, and the intersection with the red vertical line gives you the prediction of the new value.

在方框中輸入任何數值，我們就能知道根據迴歸曲線所預測的對應y 坐標值。你還可以看到迴歸線如何延伸，而與紅色垂直線的交點即為新值的預測結果。

For example, in our case, if we want to predict the average hours of daylight of March in the next year, we can type 15 into the box and we can see it will be about 12.7 hours. 例如，在我們的情況中，如果想預測明年三月的平均日照時數，可以在方框中輸入15，然後就會看到預測結果約為12.7 小時。



3 Conclusion

總結

Nonlinear regression is a very powerful tool in mathematical modelling. By considering different nonlinear regression models, we can understand the trends in datasets and make predictions. We hope that this R Shiny General Fitting (XY) tool may help you with your mathematical modelling journey. Good luck!

非線性迴歸是數學建模中非常強大的工具。通過考慮不同的非線性迴歸模型，我們可以了解數據集中的趨勢並進行預測。我們希望這個R Shiny 一般擬合(XY) 工具能夠在你的數學建模旅程中提供幫助。祝你一切順利！