*Mapana Journal of Sciences*
2020, Vol. 19, No. 3, 11-25
ISSN 0975-3303 | https://doi.org/10.12723/mjs.54.2

# A Gaussian Modelling Study of Recurring Waves of New Covid-19 Infections in Indian States

M N Anandaram* and N G Puttaswamy†

## Abstract

Countries are working very hard to control the spread of new Covid-19 infections. When the numbers of cases come down the people tend to relax controls quickly. This has resulted in a second wave of infections. European countries especially France, Italy, UK, and Germany have clearly exhibited this (bigger) second wave after noticing the low-level of infections in the earlier stages [1]. United States is now going through a third wave. In India while Delhi is ending its third wave, many states such as Haryana, Rajasthan and Madhya Pradesh have now begun showing clear signs of second wave; the possibility of a second wave is also exhibited in the case of Karnataka state. We have analyzed these trends using Gaussian model combinations and point out the need to adhere to all safety norms for many more months as we feel that more waves are still possible but unpredictable at this time.

**Keywords**: Waves of Covid-19 infections, 2-Gaussian and 3-Gaussian Model Fits, LMFIT Software, Delhi case, Karnataka case.

* Professor of Physics (Retired), Bangalore University, Bangalore, India
mnanandaram@gmail.com
† Professor of Physics (Retired), Bangalore University, Bangalore, India;

### 1.1 World Scenario

As an example, we show in Fig. 1 the daily new cases (a rolling 7-day average, per million population) for United States and four European Countries of France, Italy, UK and Germany [2].

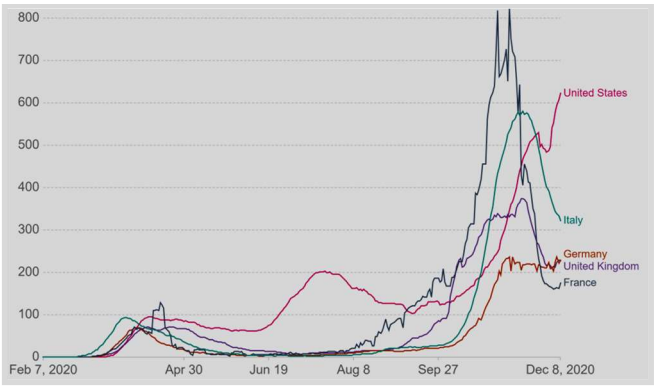



**Figure. 1:** New confirmed Covid-19 cases in the United States and some European countries [2].

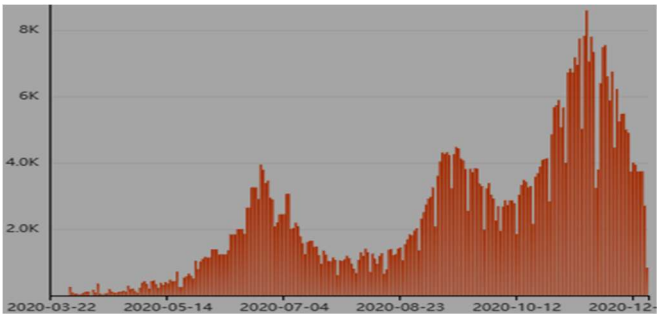
After the initial peak of cases in April, the European countries had low-level of infections during May, June, and July of 2020; however, a second wave of infections has broken out in August/September. Now the number of cases is coming down. In the United States there was a second wave in July/August and now, since September 2020 a third wave has started.

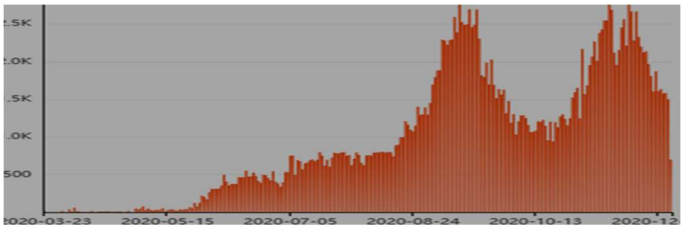
### 1.2 Scenario in Indian States

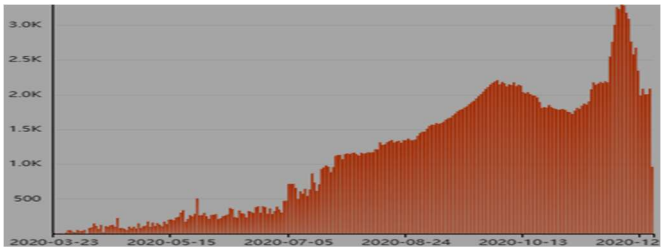
In an earlier study [3] a detailed analysis of new coronavirus infections was done for India data as a whole. In that case details of repeating waves of infection surges and decays had merged and so were not available. Hence state wise details are examined in this paper.

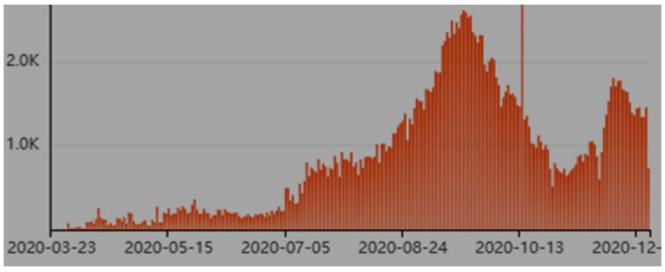
While Delhi UT [4] (hereafter referred to as Delhi) in Fig.2 (top) shows evidence of clear second and strong third waves, only some states in India have witnessed a strong second wave. Here, in Fig. 2, we also give the examples of similar waves in Haryana [5], Rajasthan [6], Madhya Pradesh [7] and Karnataka [8]. Thus the evidence for a

second wave is quite strong in the first three states except in Karnataka.



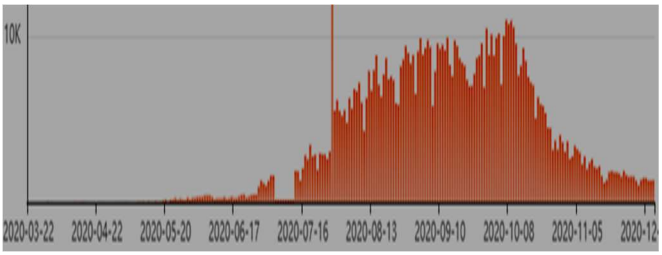



**Figure. 2:** Daily number of new confirmed Covid-19 cases in Delhi (top), Haryana (below top), Rajasthan (middle), Madhya Pradesh (lower) and Karnataka (bottom) [4, 5, 6, 7, 8].

We will first examine whether the Karnataka data contains just a single wave or a mixture of two waves by using the versatile LMFIT Python package and applying a single-Gaussian and then a two-Gaussian model based least square curve fitting analysis. Then we will apply a similar analysis to the Delhi data using a three-Gaussian model. The definition of the three Gaussian functions used in our work is provided below.

The standard definition of a Gaussian function, $\boldsymbol{f(x, A, \mu, \sigma)}$, is given by,

$$\boldsymbol{f(x, A, \mu, \sigma) = \frac{A}{\sigma\sqrt{2\pi}} exp\left(-(x-\mu)^2/2\sigma^2\right)} \quad (1.1)$$

Here $\boldsymbol{x}$ any given data point of the distribution, **A** is the amplitude, **μ** is the center (aka, mean), $\boldsymbol{\sigma}$.is the width (aka, standard deviation) and with $\boldsymbol{FWHM = 2\sigma\sqrt{2ln(2)}}$. A 2-Gaussian mixture model is formed by the addition of two 1-Gaussian functions given above but the argument parameters of each and in particular μ must not be **coincident**. The expression is given by,

$$\boldsymbol{f(x, A, \mu, \sigma) = f(x, A_1, \mu_1, \sigma_1) + f(x, A_2, \mu_2, \sigma_2)} \quad (1.2)$$

Similarly to get a 3-Gaussian mixture model we may add another term $\boldsymbol{f(x, A_3, \mu_3, \sigma_3)}$ to Eqn.(1.2). This is given, with the obvious stipulation that $\boldsymbol{\mu_1 \neq \mu_2 \neq \mu_3}$, by

$$\boldsymbol{f(x, A, \mu, \sigma) = f(x, A_1, \mu_1, \sigma_1) + f(x, A_2, \mu_2, \sigma_2) + f(x, A_3, \mu_3, \sigma_3)} \quad (1.3)$$

The LMFIT package can work out after a number of functional evaluations the minimized least square parameters of all the

component gaussians involved after starting from given set of initially estimated parameters and their allowed range limits.

## 2. The case of Karnataka Covid-19 Data

In the case of Karnataka state (Figure.2 bottom, Fig.3 and Fig. A-I in Appendix-I), the daily new confirmed cases show a broad peak overall [8], but there is a small gap at 190$^{th}$ day and, there is a hint of a second peak near 200$^{th}$ day (7$^{th}$ October).

We have investigated the data for Karnataka to see if the second wave is discernible and can be separated from the dominating first wave behind it. We first show in Figure.3 a single-Gaussian model fit to the Karnataka data set of the daily new cases of Covid-19 infection covering the period from March 22 to December 12, 2020 with, a total of 266 days or data points [8]. The LMFIT package carried out the least square minimized curve fitting of this dataset after it was provided with starting values of all Gaussian parameters and range limits of allowed values for those parameters.

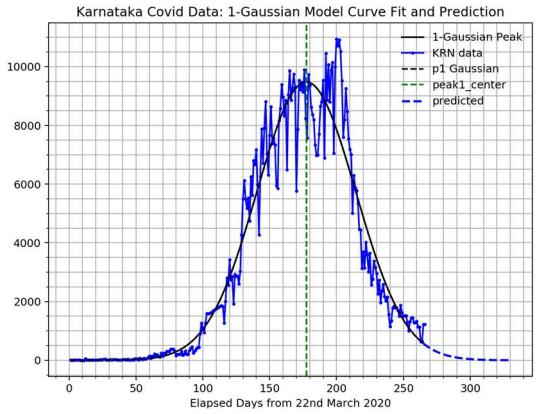



**Figure. 3:** Single-Gaussian model fit to the Karnataka Covid-19 data (linear plot). The peak position is at 177$^{th}$ day with a FWHM of 87.0 days.

As can be seen from Figure 1 and Fig. A-I (Appendix-I), a possible second peak can be suspected to lie at around 210$^{th}$ day. To see if this is real we did a LMFIT least square fit analysis with a combination of two-Gaussian models (Eqn. 1.2) on this dataset. The result

obtained after 265 evaluations fits the Karnataka data better as shown in Figure.4.

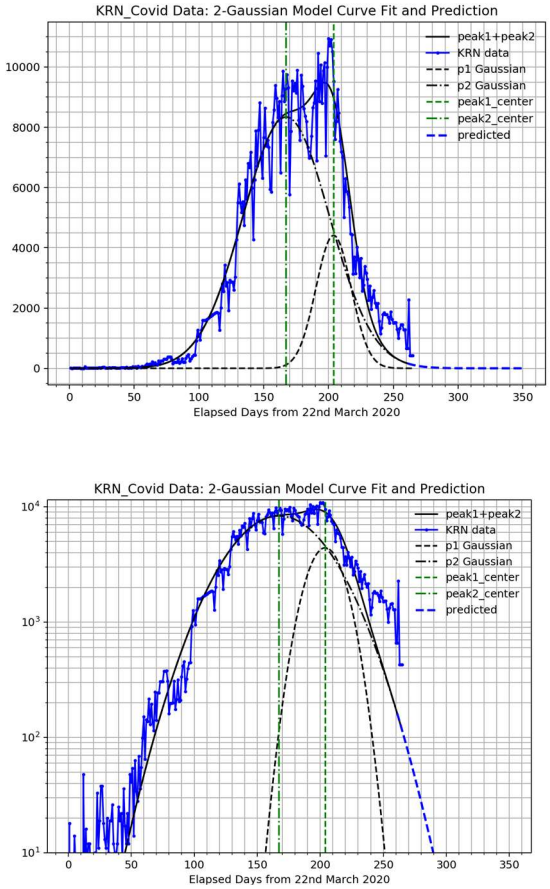



**Figure 4**: (Top: linear plot) Two-Gaussian model fit to Karnataka Covid cases . (Bottom: semilog plot) The 1st peak occurs on 167th day and 2nd peak occurs on 204th day.

The height of second peak is about 50% shorter than the first one. It is possible that the second wave could have started in some district or big city in Karnataka. However, we also note that the second wave and the first wave are both coming down quite quickly.

***A semi-log plot of Karnataka data for the single-Gaussian fit indicates that the number of new cases will come down to 100 by 4th January and to around to 10 cases by about 14th January 2021.***

In the Table 3.1 we summarize the results and provide predictions as to when the daily Covid cases in Karnataka would reach low levels of 100, 50 and 10. The single Gaussian modelled linear plot (Fig.4) has the symmetrical bell shaped profile whereas the two-Gaussian modelled plot (Fig.5) shows longer falling times for the daily new coronavirus infection cases. The best least square curve fitting parameters for the both Gaussian models are listed in Table 3.1.

The elapsed days and dates on which the new confirmed infections fall reach 100, 50 and 10 daily new case levels are listed in Table 3.2. It may be seen that for the 2-Gaussian model reaches those levels about 30 days earlier than the single-Gaussian model as shown in the second and third rows respectively in Table 3.2. Thus these two models indicate that the coronavirus infection reaches around 10 daily new infection levels by 1st week of January 2021 and 2nd week of January 2021, respectively. There is no indication yet of any third wave starting after this 2nd wave.

Table 3.1. LMFIT Parameters for Single and Double Gaussian Models applied to Karnataka Covid-19 Data

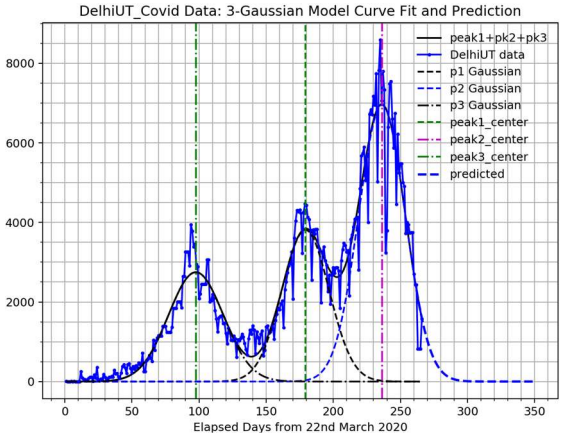
| Model | No. of Peaks | Amplitude | height | $\mu$ | $\sigma$ | fwhm |
|---|---|---|---|---|---|---|
| 1-Gausssian | Single peak | 877606.6 | 9477.7 | 177.4 | 36.9 | 87.0 |
| 2-Gausssian | 1st Peak | 697007.6 | 8339.1 | 167.3 | 31.3 | 78.5 |
| | 2nd Peak | 150000.0 | 4410.1 | 203.9 | 13.6 | 32.0 |

| Table 3.2. Name of Model | Elapsed Days from 22nd March 2020 for Covid-19 Densities ( new infection cases in Karnataka) of 100 | 50 | 10 |
|---|---|---|---|
| 2-Gaussian Dates→ | 266th day = Dec 12, 2020 | 274th day = Dec 20, 2020 | 290th day = Jan 06, 2021 |
| 1-Gaussian Dates → | 289th day = Jan 04, 2021 | 297th day = Jan 07, 2021 | 314th day = Jan 14, 2021 |

There is a possibility that the coronavirus infection may continue to exist at low levels as an endemic disease like many other viral and bacterial diseases. It is thus necessary to study actual new infection data for the next few months to get better predictions for 2021.

## 3. The case of Delhi Covid-19 Data

The Delhi dataset consists of a 266 day-data pairs starting from 22nd March until 12th Dec 2020. The plots of the Delhi Covid-19 dataset shown in Fig.2 (top) and Fig A-II (in Appendix A-II) indicates three well separated peaks near 100th day, 180th day and 240th day respectively. Hence we carried out a LMFIT least square curve fitting exercise with a combination of three-Gaussian models applied to the dataset. The Python script listed in the Appendix-II was written to accommodate the starting values and their allowed ranges of all parameters separately for each Gaussian component and after 587 function evaluations the LMFIT did the job successfully. The resulting plot is shown both as a linear plot and as a semilog plot in Figure 5 below.



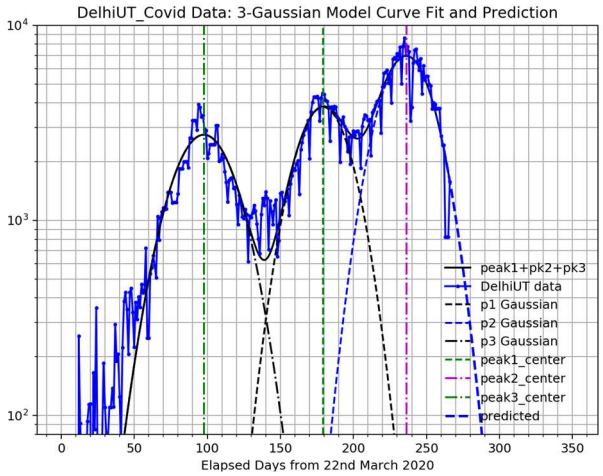



**Figure 5**. (Top) Linear plot of three-Gaussian model fit to Delhi Covid cases. The 1st peak occurs on 98th day (FWHM = 51.4), 2nd peak occurs on 179th day (FWHM = 38.5) and 3rd peak is at 236th day (FWHM = 43.2). (Below) Semilog plot). All curves have the typical Gaussian bell shape.

The semilog plot shows that the third wave reaches very low levels of new infections before 2nd week of January 2021. There is no indication yet about any start of a fourth wave surge.

## 4. Conclusions

We were able to successfully apply single-Gaussian and two-Gaussian combination models from the convenient LMFIT package to the Karnataka Covid-19 dataset. However, only the three-Gaussian combination model was applied to the Delhi dataset as it had three peaks. Similar work can be done with other relevant Covid-19 datasets. Both the two-Gaussian and the three-Gaussian results confirm the bell shape of all Gaussian curves and also indicate that the corresponding Covid-19 waves undergo quick surges and drawdowns to very low levels of the new infection cases. However there were no indications as to whether or not another wave surge will soon follow the lowest points. That has to await further data collection in the next few months. Since a possibility of another wave surge cannot be ruled out all precautions must continue to be maintained.

## References


[1] https://www.nature.com/articles/s41598-020-72611-5 September 23, 2020

[2] https://ourworldindata.org/coronavirus-data-explorer?zoomToSelection=true&time=2020-02-07..latest&country=USA~GBR~DEU~FRA~ITA&region=World&casesMetric=true&interval=smoothed&hideControls=true&perCapita=true&smoothing=7&pickerMetric=total_cases&pickerSort=desc

[3] N.G. Puttaswamy and M.N. Anandaram, "Current Trends Indicate that the COVID-19 Curve Has Flattened in India: The daily New Cases Are Expected to Reach a Low Value by End of February 2021", Preprint of Paper submitted to MJOS

[4] https://www.bing.com/covid/local/delhi_india?vert=graph

[5] https://www.bing.com/covid/local/haryana_india?vert=graph

[6] https://www.bing.com/covid/local/rajasthan_india?vert=graph

[7] https://www.bing.com/covid/local/madhyapradesh_india?vert=graph

[8] https://www.bing.com/covid/local/karnataka_india?vert=graph

[9] https://lmfit.github.io/lmfit-py/


## Appendix-I. Karnataka Covid Data Plotting Code Listing (Python 3.7)

The Covid wave data listings for Karnataka given here and for Delhi in the next Appendix-II are provided to allow interested readers to use them in their own analysis. They can also get more data beyond 12th December by linking to [4] and/or [8] and read off the plot therein.

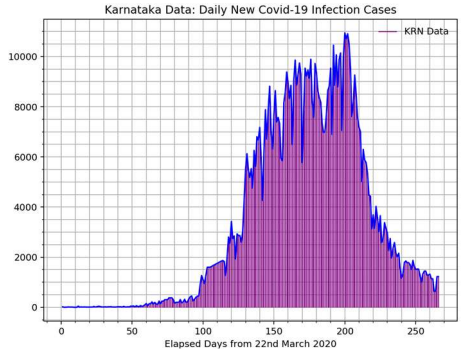



**Figure A-I**: Plot of Karnataka Data done by the listing below

This Listed Karnataka Covid data below produces a plot shown above

```
import numpy as np
import matplotlib.pyplot as plt
#import lmfit
# KRN COVID-19 X , Y data from 22nd March to 12th Dec 2020 (266 days)
#Source:
https://www.bing.com/covid/local/karnataka_india?vert=graph
x = np.linspace( 1.0, 266.0, 266);
y = np.array([
 18., 0, 4, 4, 14, 7, 7, 7, 0, 0, 0, 48, 4, 16, 7, 12,
 12, 6, 8, 8, 10, 7, 33, 11, 19, 38, 38, 18, 13, 11, 20, 10,
 19, 19, 26, 12, 10, 9, 12, 25, 19, 22, 8, 36, 17, 12, 22,
 12, 48, 41, 54, 14, 63, 34, 28, 69, 36, 55, 99, 151, 65,
 143, 138, 216, 130, 194, 135, 115, 248, 141, 243, 243, 304,
 304, 304, 377, 377, 378, 308, 161, 198, 198, 199, 308, 176,
 213, 317, 204, 210, 337, 416, 453, 249, 322, 397, 442, 445,918,
1267, 1105, 947, 1272, 1598, 1598, 1598, 1629, 1660, 1691, 1722,
1753,
1784, 1815, 1846, 1863, 1826, 1269, 2014, 2802, 2550, 3426, 2739,
2846, 1924, 2929, 2862, 2862, 2597, 3029, 4267, 5503, 6128, 5483,
5172, 5532, 4752, 6259, 5619, 6805, 6670, 7178, 5985, 4267, 6257,
7883, 6706, 7908, 8818, 7040, 6317, 7665, 8642, 7385, 7571, 7330,
5938,
5851, 8161, 8580, 9386, 8960, 8324, 8852, 6495, 9058, 9860, 8865,
9280,
9746, 9319, 5773, 7866, 9540, 9217, 9464, 9140, 9894, 8244, 7576,
9725,
9366, 8626, 8364, 8191, 7339, 6974, 6997, 7710, 8655, 8811, 9543,
6892,
10453, 8856, 10070, 8793, 9886, 10145, 7051, 9993, 10947, 10704,
10913, 10517, 9523, 7606, 8191, 9265, 8477, 7542, 7184, 7012, 5018,
6297, 5872, 5778, 5356, 4471, 4439, 3130, 3691, 3146, 4025, 3589,
3014,
3652, 2576, 2756, 3377, 3156, 2960, 2258, 2740, 1963, 2362, 2584,
2166,
2016, 2154, 1565, 1157, 1336, 1791, 1849, 1781, 1781, 1704, 1509,
1870,
```

```
1630, 1505, 1526, 1522, 1291, 998.0,
1330, 1440,1446, 1286, 1286, 1321,1139,1139, 639,640,1224, 1224] )
#for ii in range (len(y)): print(ii,"-->", x[ii],"-->",y[ii])
plt.figure( figsize = (8,6), dpi = 180)
plt.vlines( x, ymin=0, ymax=y, colors='purple', lw=1 )
plt.plot(x, y, "b", label = 'KRN Data' )
plt.grid( which = "both" ); plt.minorticks_on()
plt.xlabel("Elapsed Days from 22nd March 2020")
plt.legend( loc= “best”, frameon = False)
plt.title("Karnataka Data: Daily New Covid-19 Infection Cases")
plt.show()
# for LMFIT Gaussian Fitting refer to code in Appendix-II given
next #
```

## Appendix-II. Delhi Covid Data Analysis Code Listing (Python 3.7)

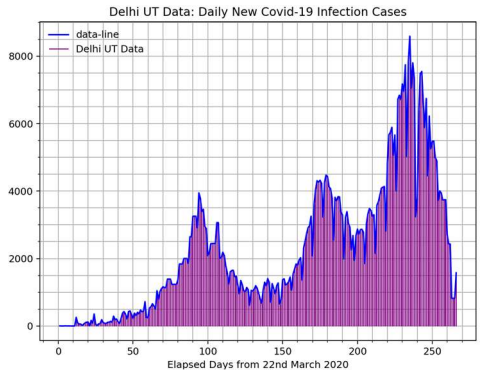



**Figure. A-II**. This plot of the Delhi [4] Covid data can be produced by the listed y-data in the code listing given below.

Python code for a 3-Gaussian Model Least Square Fit to the Delhi Covid Data

```
# -*- coding: utf-8 -*- # This Script was written by Mandyam
N. Anandaram
import numpy as np
Import matplotlib.pyplot as plt
import lmfit
# KRN COVID-19 X and Y data from 22nd March to 12th Dec
2020 (266 days)
```

```
#Source:
https://www.bing.com/covid/local/delhi_india?vert=graph
x = np.linspace( 1.0, 266.0, 266 );    # 266 days to 12th Dec 2020
y = np.array([ 3, 0,  1,  1,  5,  0,  3,  0,  0,  0,
0,  256,  91,  59,  58,  20,  53,  93,  114,  115,
5,  166,  85,  356,  51,  17,  62,  67,  186,  110,
78,  75,  110,  110,  138,  111,  293,  190,  206,  125,
76,  223,  384,  427,  349,  206,  428,  448,  338,  224,
381,  310,  406,  359,  472,  425,  438,  721,  250,  250,
534,  571,  660,  591,  508, 1047,  792, 1024, 1105, 1163, 1142,
1143, 1390, 1390, 1390, 1234, 1235, 1235, 1235, 1366, 1838,
1838, 1839, 2001, 2002, 2002, 1859, 2645, 2646, 3255, 3256,
3256, 2909, 3947, 3788, 3390, 3460, 2948, 2889, 2084, 2199,
2442, 2446, 2446, 2447, 3064, 3064, 2008, 2033, 2187, 2089,
1781, 1573, 1246, 1606, 1647, 1652, 1462, 1475, 1211, 954,
1349, 1227, 1041, 1025, 1142, 1075,  613, 1056, 1035, 1093, 1195,
1118,  961,  805,  674, 1076, 1299, 1192, 1404, 1300, 707,
1257, 1113,  956, 1192, 1276,  652,  787, 1374, 1398, 1215,
1250, 1312, 1450, 1061, 1544, 1693, 1840, 1808, 1954, 2024, 1358,
2312, 2509, 2737, 2914, 2973, 3256, 2077, 3609, 4039, 4308,
4266, 4321, 4235, 3229, 4263, 4473, 4432, 4127, 4071, 3812,
2548, 3816, 3714, 3834, 3827, 3372, 3292, 1984, 3227, 3390,
3037, 2920, 2258, 2683, 1947, 2676, 2871, 2726, 2860, 2866,
2780, 1849, 3036, 3324, 3483, 3428, 3259, 3299, 2154, 3579,
3686, 3882, 4086, 4116, 4136, 2812, 4853, 5673, 5739, 5891, 5062,
5664, 4001, 6725, 6842, 6715, 7178, 6953, 7745, 5023, 7830,
8593, 7053, 7802, 7340, 3235, 3797, 6396, 7486, 7546, 6608,
5879, 6746, 4454, 6224, 5246, 5475, 5482, 4998, 4906, 3726,
4006, 3944, 3734,3743, 3743, 2706, 2431, 2431,821,821,
821. , 1575] )
#for ii in range (len(y)): print(ii,"-->", x[ii],"-->",y[ii])
plt.figure( figsize = (8,6), dpi=180)   # plot covid data-set
plt.vlines(x,ymin=0,ymax=y,colors='purple',lw=1,label='Delhi
UT Data')
plt.plot( x, y, "b", label = "data-line")
plt.grid( which = "both" );    plt.minorticks_on()
plt.xlabel("Elapsed Days from 22nd March 2020")
plt.legend( loc=0, frameon = False )
plt.title("Delhi UT Data: Daily New Covid-19 Infection Cases")
plt.show()
```

```
# ## Model definition and Gaussian component definitions
peak1 = lmfit.models.GaussianModel(prefix='peak1_')
peak2 = lmfit.models.GaussianModel(prefix='peak2_')
peak3 = lmfit.models.GaussianModel(prefix='peak3_')
model = peak1 + peak2 + peak3
model.set_param_hint('peak1_center', value=170., min=80,
max=260)
model.set_param_hint('peak2_center', value=185., min=85,
max=270)
model.set_param_hint('peak3_center', value=200., min=90,
max=280)
model.set_param_hint('peak1_sigma', value=30., min=10.0,
max=50.0)
model.set_param_hint('peak2_sigma', value=40., min=10.0,
max=70.0)
model.set_param_hint('peak3_sigma', value=50., min=10.0,
max=80.0)
model.set_param_hint('peak1_amplitude', value=7.0e5,
min=1e5,max=1e6)
model.set_param_hint('peak2_amplitude', value=7.5e5,
min=1e5,max=1e6)
model.set_param_hint('peak3_amplitude', value=7.7e5,
min=1e5,max=1e6)

name = '3-gaussians'
#  Carry out the double-Gaussian model Fit
fit_res = model.fit(y, x=x)
print (fit_res.fit_report()) # output the model fitment report
x_pred = np.linspace(266.0,350.0, 89)  # points for
extrapolation
y_pred = fit_res.eval(x=x_pred)  # predicted extrapolation

# Graph all results in a plot
plt.subplots( figsize = (8,6), dpi = 180)
x = x
plt.plot(x, model.eval(x=x, **fit_res.values),'k-
',label="peak1+pk2+pk3")
plt.plot(x, y, 'bo-', ms=2, label="DelhiUT data");
p1p2p3 = [ 'p1 Gaussian', ‘p2 Gaussian', 'p3 Gaussian' ]
if fit_res.model.components is not None:
```

```
for component, p123 in zip(fit_res.model.components,
p1p2p3):
        if p123=='p1 Gaussian':
           plt.plot(x, component.eval(x=x,**fit_res.values),'k--
',label=p123)
        if p123=='p2 Gaussian':
           plt.plot(x, component.eval(x=x,**fit_res.values),'b--
',label=p123)
        if p123=='p3 Gaussian':
           plt.plot(x, component.eval(x=x,**fit_res.values),'k-
.',label=p123)
for param in ['peak1_center', 'peak2_center', 'peak3_center']:
    if param == 'peak1_center':
       plt.axvline(fit_res.params[param].value,ls='--',color='g',
label=param)
    if param == 'peak2_center':
       plt.axvline(fit_res.params[param].value,ls='-.',color='m',
label=param)
    if param == 'peak3_center':
       plt.axvline(fit_res.params[param].value,ls='-.',color='g',
label=param)

plt.plot(x_pred,y_pred,"b--",lw=2,label="predicted")
plt.grid(which="both");   plt.minorticks_on();
plt.yscale("log");  plt.ylim( [80,  1.0e4] ) #← uncomment to get
Log(Y) plot
plt.xlabel("Elapsed Days from 22nd March 2020")
plt.legend(loc="best",frameon=False)
plt.title("DelhiUT_Covid Data: 3-Gaussian Model Curve Fit
and Prediction")
plt.show()
```