Mathematical Modelling and Analysis I

Coursework 2

MODEL 1:

(a) Deriving a mathematical model for the amount of pollutant in the lake as a function of time i.e., y(t)

We must find the input and output rates of pollutant.

Input rate is obtained as the product of the concentration of pollutant and the rate of flow of water

Input Rate = c_i r_i

For the output rate we must first find the concentration of pollutant of outflow which is the same as the concentration of pollutant.

Using the relation
Concentration (density) = Mass/Volume
We get,
Outflow concentration (c_o) = y(t)/V(t)

Output Rate = y(t).r_o / V(t)

However, in order to simplify calculation, we will assume the volume to be constant so V(t) = V

The ODE "dv/dt = r_i – r_o" shows how the lakes volume changes with respect to time. This may lead to three outcomes

1) If $r_i > r_o$ the lakes volume increases linearly with time
2) If $r_i = r_o$ the volume remains constant with time
3) If $r_i < r_o$ the volume decreases linearly with time

Since these cases all depend on various factors, we are not aware of such as the size of the outlet, its orientation etc. we are forced to assume one of these outcomes is true. For the simplification of calculations, we can assume case 2) that is the volume of the lake remains constant with time i.e., $r_i = r_o$ and V(t) = V => constant

Now, change in amount of pollutant is given by

dy/dt = input rate – output rate

⇨ $dy/dt = c_i r_i - y(t).r_o / V$

⇨ $dy/dt + y(t).r_o / V = c_i r_i$ ....(ODE 1)

Solving the above differential equation using integrating factor (IF) method
Because it is of the form

dy/dt + y. P(t) = Q(t)

Where, $P(t) = r_0 / V$, $Q(t) = c_i r_i$

$IF = e^{\int P(t)\, dt}$

⇨ $IF = e^{\int \frac{r_o}{V} dt}$

⇨ $\mathbf{IF = e^{\frac{r_o}{V}t}}$ (r_o / V is a constant term)

Now multiplying (ODE 1) by IF

⇨ $dy/dt \,.\, (e^{\frac{r_o}{V}t}) + y(t).r_o (e^{\frac{r_o}{V}t}) / V = c_i r_i (e^{\frac{r_o}{V}t})$

⇨ $\frac{d}{dt}(y(t) \,.\, e^{\frac{r_o}{V}t}) = c_i r_i (e^{\frac{r_o}{V}t})$

> Using product rule of differentiation, we have:
> $\frac{d}{dt} U.V = V.\frac{du}{dt} + U.\frac{dv}{dt}$

Integrating both sides

⇨ $y(t)\, e^{\frac{r_o}{V}t} = \int (c_i r_i \,.\, e^{\frac{r_o}{V}t})\, dt + K$ (where K is the constant of integration)

⇨ $y(t)\, e^{\frac{r_o}{V}t} = c_i r_i \int e^{\frac{r_o}{V}t}\, dt + K$ ($c_i r_i$ is a constant term)

⇨ $y(t)\, e^{\frac{r_o}{V}t} = c_i r_i [(V/r_o).\ e^{\frac{r_o}{V}t}] + K$ ($\int e^x\, dx = e^x$, r_o / V is a constant term so we divide by it after integrating)

⇨ $y(t)\, e^{\frac{r_o}{V}t} = c_i r_i [(V/r_i).\, e^{\frac{r_o}{V}t}] + K$ (from assumption we have $r_i = r_o$)

⇨ $y(t)\, e^{\frac{r_o}{V}t} = c_i V.\ e^{\frac{r_o}{V}t} + K$

⇨ $y(t) = c_i V + K/e^{\frac{r_o}{V}t}$ (divide equation by $e^{\frac{r_o}{V}t}$)

⇨ $y(t) = c_i V + K/e^{\frac{r_o}{V}t}$

⇨ $\mathbf{y(t) = K.e^{\frac{-r_o}{V}t} + c_i V}$ ...... (i)

We know that at t = 0, y(t) = 0
Plugging these values into eq (i) we get
$K = - c_i V$

⇨ $y(t) = c_i V - c_i V.\ e^{\frac{-r_o}{V}t}$

⇨ $\boxed{\mathbf{y(t) = c_i V (1 - e^{\frac{-r_o}{V}t})\ g}}$ ...... (ii)

DIMENSIONAL CONSISTENCY OF THE MODEL

For **y(t)**
It is the amount of pollutant in the lake and is measured in g so has the dimensions [M]

For $\mathbf{c_i}$
It is concentration of pollutant and is measured in g/m^3 so has dimensions $[M\ L^{-3}]$

For **V**
It is volume so has dimensions $[L^3]$

For $\mathbf{(1-e^{\frac{-r_o}{V}t})}$
1 is a constant so is dimensionless and all exponential functions are also always dimensionless, so this term has no dimensions

LHS:
y(t)
⇨ [M]

RHS:
$c_iV\ (1- e^{\frac{-r_o}{V}t})$
⇨ $[M\ L^{-3}].[L^3]$
⇨ [M]

LHS = RHS
Thus, this model is dimensionally correct.

(b) Using the equation (ii) from (a) we derive an expression for the time varying with concentration

We have,
$y(t) = c_iV\ (1- e^{\frac{-r_o}{V}t})$

⇨ $y(t)/c_iV = (1- e^{\frac{-r_o}{V}t})$

⇨ $e^{\frac{-r_o}{V}t} = 1 - (y(t)/c_iV)$

Taking logarithm of both sides

⇨ $\ln(e^{\frac{-r_o}{V}t}) = \ln(1 - (y(t)/c_iV))$ $(\ln(e^x) = x)$

⇨ $-r_0\ t/ V = \ln(1 - (y(t)/c_iV))$

⇨ $\boxed{\mathbf{t = \frac{-V}{r_o} \ln(1 - (y(t)/c_iV))\ hours}}$ ...... (iii)

We are given that the lake will be irreversibly damaged when its concentration reaches $331\ g\ m^{-3}$
In order to find the time of irreversible damage we should estimate the values of the constants V, r_i , r_o and c_i
Considering a case study of Lake Pichola in Rajasthan, India

We get
$V = 13.08 \times 10^6\ m^3$ [1]
$r_i = r_o = 45000\ m^3/h$
In order to ensure that c_i is of similar order of c_x we assume,
$c_i = 450\ g/m^3$
Also we write $y(t)/V = c_o$ (concentration of outflow)

So, the equation (ii) becomes

$$c_o = c_i\,(1 - e^{\frac{-r_o}{V}t}) \qquad ....(iv)$$

And, the equation (iii) becomes

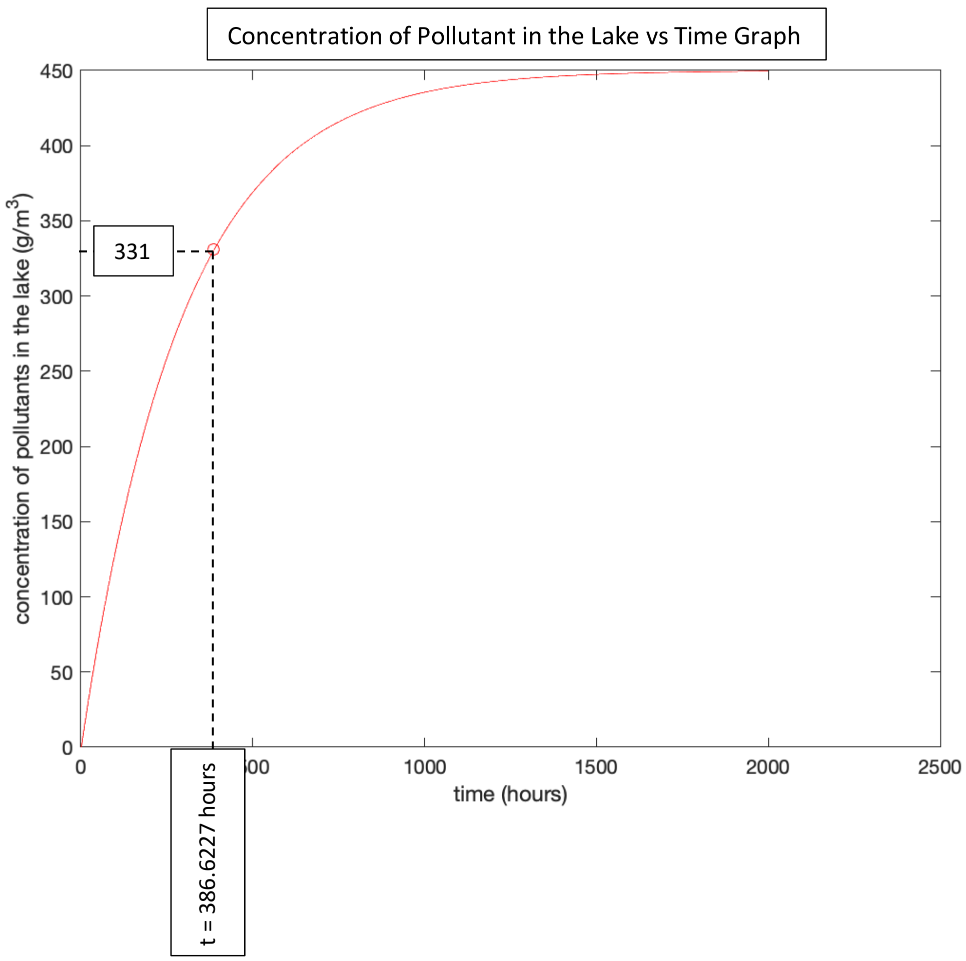
$$t = \frac{-V}{r_o}\ln(1 - (c_0/c_i)) \qquad ......(v)$$

For the irreversible damage it is given that the outflow concentration must be $331\ g/m^3$
Let this be called the critical concentration represented by c_x
$c_x = 331\ g/m^3$

We plug the assumed values into equation (v) where c_x replaces c_o
On solving through MatLab we get the value of t [APPENDIX 1]

t = 386.6227 hours

We graph equation (iv) using Matlab and plot the value of 't' obtained on it to get the resultant graph [APPENDIX 2]



(c) First, we start by calculating he steady state concentration of the lake given by

$$\lim_{t\to\infty} \frac{y\,(t)}{V}$$

We have y(t) = c_iV (1 - $e^{\frac{-r_o}{V}t}$)

⇨ y(t) / V = c_i - c_i $e^{\frac{-r_o}{V}t}$

⇨ $\lim_{t\to\infty} \frac{y\,(t)}{V} = \lim_{t\to\infty} (c_i - c_i\, e^{\frac{-r_o}{V}t})$

⇨ $\lim_{t\to\infty} \frac{y\,(t)}{V} = \lim_{t\to\infty} c_i - \lim_{t\to\infty} c_i\, e^{\frac{-r_o}{V}t}$

> Calculating $\lim_{t\to\infty} c_i\, e^{\frac{-r_o}{V}t}$
>
> ⇨ $c_i \lim_{t\to\infty} e^{\frac{-r_o}{V}t}$
>
> We know that $\lim_{x\to\infty} e^{-x} = 0$
>
> ⇨ 0

⇨ $\lim_{t\to\infty} \frac{y\,(t)}{V} = \lim_{t\to\infty} c_i$

We have chosen c_i to be a constant with c_i = 450 g/m^3

⇨ $\lim_{t\to\infty} \frac{y\,(t)}{V} = c_i$

⇨ $\lim_{t\to\infty} \frac{y\,(t)}{V} = 450$ (this is the steady state concentration)

It is given that at time τ, concentration reaches 95% of steady state concentration

Thus, at t = τ, c_L (τ) = 95% of 450

⇨ c_L (τ) = 427.5 g/m^3

We have from equation (v)

t = $\frac{-V}{r_o}$ ln(1 –(c_0/c_i))

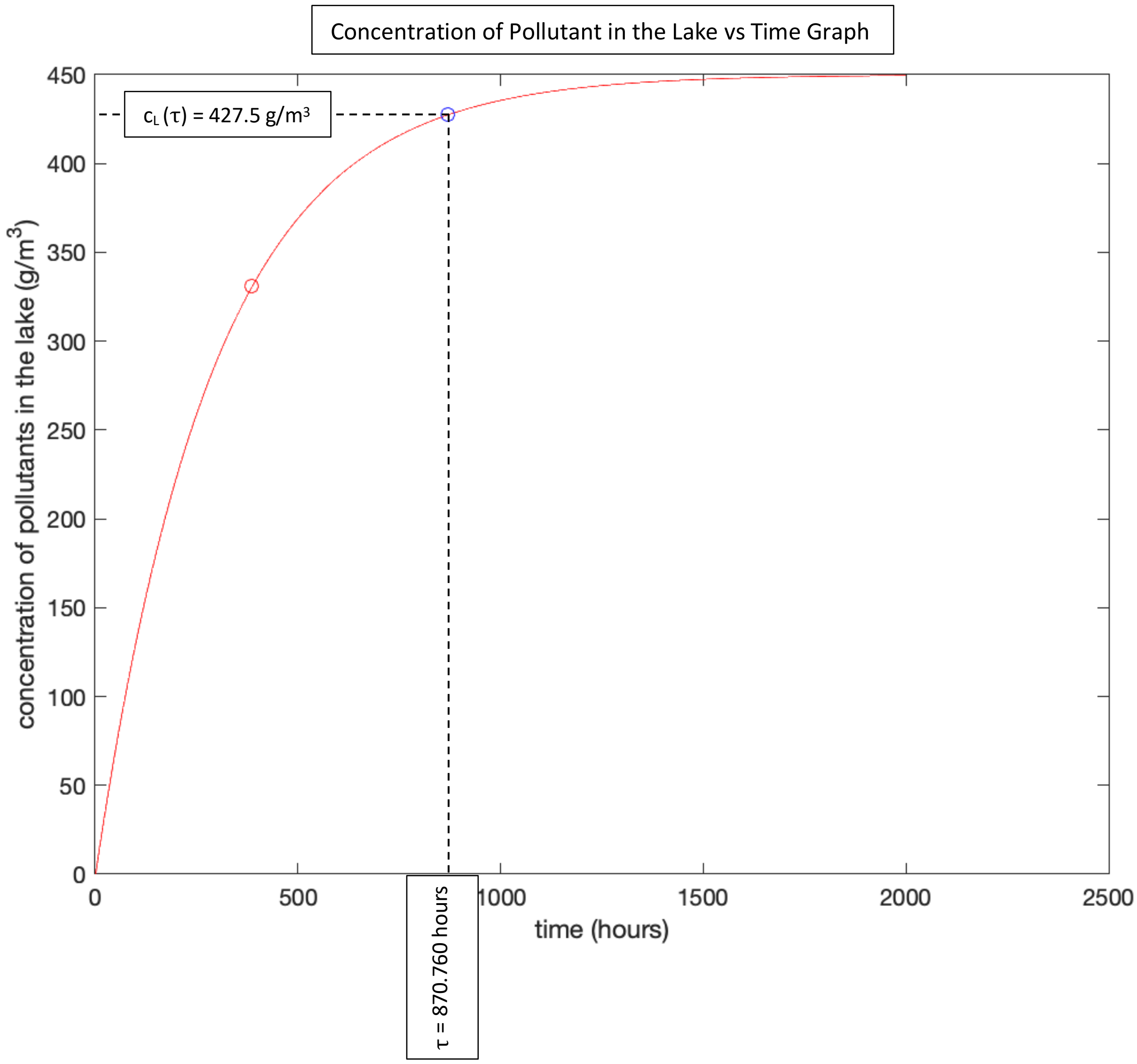
⇨ τ = $\frac{-V}{r_o}$ ln(1 – (c_L (τ) /c_i))

Using values of V, r_o and c_i as specified in (b) [APPENDIX 3]

⇨ τ = $\frac{-13.08 \times 10^6}{45000}$ ln(1 – (427.5 /450))

⇨ **τ = 870.760 hours**

GRAPH SHOWING TIME τ AND REGION OF STEADY STATE



In the above plot showing a plot of concentration of pollutants in the lake with respect to time, we can observe that the graph becomes a line parallel to the x axis when it approaches the concentration 450 g/m^3. Thus, the **region of steady state concentration** is the region of time = ∞ i.e., the horizontal asymptote of the graph.

A steady state regime is when the concentration of pollutants in the lake reaches a point where it stops changing and becomes constant. This is a state wherein the concentration of pollutants in the lake reaches equilibrium. However, since this is only a hypothetical situation as it's a horizontal asymptote so it never really reaches steady state, the steady state regime can be defined as a state of hypothetical equilibrium of concentration.

Since we assumed the volume of the lake to be constant, we need to consider the effect of the changing values of inflow and outflow into the lake, on the concentration of pollutant in the lake.

The change in volume of the lake is defined as follows:
$dV/dt = r_i - r_0$
⇨ $dV = (r_i - r_0)\, dt$

To determine V(t) we integrate this equation:
$\int dV = \int (r_i - r_0)\, dt$
⇨ $V = (r_i - r_0)\, t + K \;\; m^3$ (K = constant of integration)

However, at t = 0 the Volume in the lake is constant say $V = V_o$
⇨ $K = V_o$
⇨ $V(t) = (r_i - r_0)\, t + V_o \;\; m^3$

The effect of inflow and outflow into the lake can be studied by considering two main cases

When $r_i >> r_o$
⇨ $(r_i - r_0)\, t > 0$
⇨ **V(t) increases with time**
In the case where the inflow of pollutants into the lake is significantly more than the outflow, i.e., the lakes volume increase with time, the concentration of the lake reaches steady state at a slower rate. This is due to the fact that the mass of pollutant in the inflow is lower than the volume of the lake making the concentration of pollutant in the lake low. Also, the volume of the lake progressively increases with time which is another factor that contributes to the fact that in this situation it will take more time to reach steady state concentration.

When $r_i << r_o$
⇨ $(r_i - r_0)\, t < 0$
⇨ **V(t) decreases with time**
In the case where the outflow of pollutants into the lake is significantly more than the inflow, i.e., the lakes volume decreases with time, the concentration of the lake reaches steady state at a faster rate. This is due to the fact that as pollutant inflows into the lake, the pure water flows out of the lake at a faster rate. Also the volume of the lake progressively decreases with time which is another factor that contributes to the fact that in this situation it will take less time to reach steady state concentration.

It can be observed that in the above to cases as well as the case of constant volume, the concentration of the lake does eventually reach steady state concentration. The only factor that varies is the time that it takes to reach that concentration.

Also, this model cannot completely be relied on as the volume also changes due to other environmental factors such as leakage of lake water into the soil, evaporation, rain etc.

(d) In order for the lake to clear out, we need to determine a model so that it reaches a pollutant concentration less than 5 g/m^3 after the inflow of pollutant into the lake has stopped.

From (a) it was determined that the change in amount of pollutant is given by

⇨ $dy/dt = c_i\, r_i – y(t).r_o / V$

However, since the inflow into the lake has stopped we have $c_i\, r_i = 0$

⇨ $dy/dt = – y(t).r_o / V$

⇨ $dy/dt + y(t).r_o / V = 0$ ....(ODE 2)

Solving the above differential equation using integrating factor (IF) method
Because it is of the form

dy/dt + y. P(t) = Q(t)

Where, $P(t) = r_0 / V$, $Q(t) = 0$

$IF = e^{\int P(t)\, dt}$

⇨ $IF = e^{\int \frac{r_o}{V} dt}$

⇨ $\mathbf{IF = e^{\frac{r_o}{V}t}}$ (r_o / V is a constant term)

Now multiplying (ODE 2) by IF

⇨ $dy/dt \,.\, (e^{\frac{r_o}{V}t}) + y(t).r_o \,.\, (e^{\frac{r_o}{V}t}) / V = 0$

⇨ $\frac{d}{dt}(y(t) \,.\, e^{\frac{r_o}{V}t}) = 0$

⇨ $d (y(t) \,.\, e^{\frac{r_o}{V}t}) = 0.dt$

Using product rule of differentiation, we have:
$\frac{d}{dt} U.V = V.\frac{du}{dt} + U.\frac{dv}{dt}$

Integrating both sides

⇨ $y(t) \,.\, e^{\frac{r_o}{V}t} = \int 0.dt$

⇨ $y(t) \,.\, e^{\frac{r_o}{V}t} = K$ (where K is the constant of integration)

⇨ $y(t) = K \,.\, e^{\frac{-r_o}{V}t}$ ....(vi)

We know that at t = 0 (i.e., the time when inflow of pollutant stops), we know that the concentration of pollutant in the lake, i.e., $y(t)/V = 213\ g/m^3$

Plugging these values in equation (vi)

K = 213 . V

⇨ $\boxed{\mathbf{y(t) = 213.V.\, e^{\frac{-r_o}{V}t}\ g}}$ ....(vii)

Now, using equation (vii) in order to find an expression for time 't'

Rearranging equation (vii) we have

$y(t) / (213.V) = e^{\frac{-r_o}{V}t}$

Taking logarithm of both sides we have

⇨ $\ln (y(t) / (213.V)) = - r_o t / V$ $(\ln(e^x) = x)$

⇨ $\mathbf{t = \frac{-V}{r_o} \ln (y(t) / (213.V))}$ ...(viii)

Since we need to find the time when the concentration of the lake reached 5 g/m^3
Thus, $y(t)/V = 5\ g/m^3$

Replacing this value of y(t)/V and the values of V and r_o chosen in (b) in equation (viii) we get the time required to reach concentration of 5 g/m^3 [APPENDIX 4]

⇨ t = 1090.5 hours

In order for the lake to be considered cleared the concentration should be less the 5 g/m^3 so we increase this time t by a little amount and plug it into the equation [(vii)/V] in order to get a value just below 5 g/m^3

Using MatLab we get [APPENDIX 5]

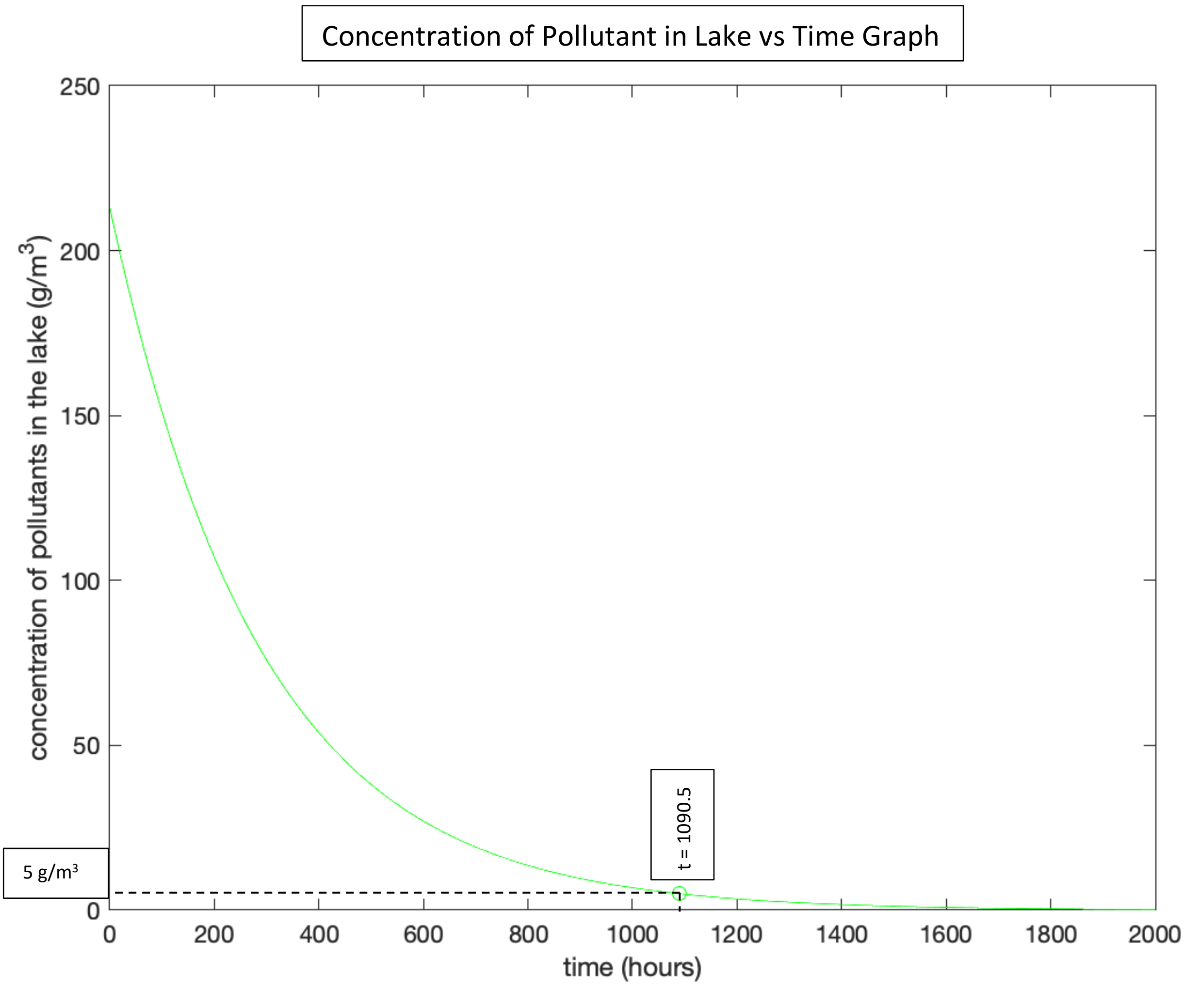
⇨ **t = 1090.55 hours**

We check the accuracy of this solution for time by plugging in this value for time back into equation [(vii)/V]

We get $y(t)/V = 4.9998\ g/m^3$

This is demonstrated in the following graph on the next page

GRAPH REPRESENTING CLEARING OF LAKE



MODEL 2:

(a) We are given the decay rate of the levels of C^{14} by the equation

$$\frac{d\,[C^{14}]}{dt} = -r\,[C^{14}]$$

Rearranging this equation, we get

$$\frac{d\,[C^{14}]}{[C^{14}]} = -r\,dt$$

Now we integrate the equation on both the sides

$$\int \frac{d\,[C^{14}]}{[C^{14}]} = -r \int dt$$

Using $\int \frac{dx}{x} = \ln x$

⇨ $\ln [C^{14}] = -rt + K$ ..... (i) (where K = constant of integration)

Finding the constant of integration K:
We consider the initial condition
We are given at $t = 0$, $[C^{14}] = [C^{14}]_0$
We plug in these values of t and $[C^{14}]$ in eq (i)

⇨ $\ln [C^{14}]_0 = -r(0) + K$

⇨ $K = \ln [C^{14}]_0$

Replacing value of K in eq (i)

⇨ $\ln [C^{14}] = -rt + \ln [C^{14}]_0$

⇨ $rt = \ln [C^{14}]_0 - \ln [C^{14}]$

⇨ $rt = \ln ([C^{14}]_0 / [C^{14}])$

⇨ $\mathbf{r = \frac{1}{t} \ln \frac{[C^{14}]_0}{[C^{14}]}\ year^{-1}}$ ....(ii)

We are given that the concentration is halved of the initial when t = 5730 years
Thus, under this condition, $[C^{14}] = 0.5\,[C^{14}]_0$
We plug these values of t and $[C^{14}]$ in equation (ii)

⇨ $r = \frac{1}{5730} \ln \frac{[C^{14}]_0}{0.5\,[C^{14}]_0}$

⇨ $r = \ln 2 / 5730$

⇨ $\mathbf{r = 1.2097 \times 10^{-4}\ year^{-1}}$

(b) We are given a relation

$M = [C^{14}]/[C^{12}]$ ....(iii) (where $[C^{12}]$ is a constant)

Now we differentiate equation (iii) with respect to 't'

⇨ $\frac{dM}{dt} = \frac{d}{dt}\left(\frac{[C^{14}]}{[C^{12}]}\right)$

⇨ $\frac{dM}{dt} = \frac{1}{[C^{12}]}\frac{d[C^{14}]}{dt}$ ....(iv) (because $[C^{12}]$ is a constant)

We have also been given

$$\frac{d\,[C^{14}]}{dt} = -r\,[C^{14}]$$

so, we replace the value of $d[C^{14}]/dt$ in equation (iv)

⇨ $\frac{dM}{dt} = \frac{1}{[C^{12}]}(-r\,[C^{14}])$

⇨ $\frac{dM}{dt} = -r\left(\frac{[C^{14}]}{[C^{12}]}\right)$

⇨ $\frac{dM}{dt} = -r\,M$ (using equation (iii))

Rearranging the equation

⇨ $\frac{dM}{M} = -\,r\,dt$

Integrating both sides of the equation

⇨ $\int \frac{dM}{M} = -\,r \int dt$ (r is a constant term)

⇨ $\ln M = -r\,t + K$ ....(v)

Using $\int \frac{dx}{x} = \ln x$

Finding the constant of integration K:
We consider the initial condition
We are given at t = 0, M = 1
We plug in these values of t and M in equation (v)

⇨ $\ln(1) = -r(0) + K$

⇨ $K = 0$ ($\ln(1) = 0$)

Replacing Value of K in equation (v)

⇨ $\ln M = -r\,t$

Taking the exponential of both sides of the equation

⇨ $\boxed{\mathbf{M = e^{-rt}}}$ ....(vi) ($e^{\ln(x)} = x$)

(c) We are given a data containing five measurements of M. Using that data we first calculate the mean and standard deviation of the data.

We have N = Sample size = 5, and X_i = different values of M from the sample

In order to calculate the **mean** we use the expression
$\overline{X} = \frac{1}{N}\Sigma(X_i)$

⇨ $\overline{X} = 0.0149$ [APPENDIX 6]

Using an **unbiased estimate for the true variance** of X, the sample variance is given by
$\mathrm{Var}(X) \approx \sigma^2 = \frac{1}{N-1}[\,(\sum_{i=1}^{N} X_i) - N\overline{X}^2]$ [APPENDIX 7]

The **standard deviation** of the sample is given by
$SD(\overline{X}) = \sqrt{\frac{\sigma^2}{N}}$

⇨ $SD(\overline{X}) = 8.9028 \times 10^{-4}$ [APPENDIX 8]

We calculate the 95% (1-α) confidence interval (x_1 , x_2) such that
$P(x_1 < \overline{X} \leq x_2) = 0.95$

Because of the symmetry of the standard normal distribution around 0
⇨ $x_1 = -x_2 = \Phi^{-1}(1-\alpha/2)$
where $\Phi^{-1}(x)$ is the inverse of the normal cumulative distributed function

By standardizing the random variable $\overline{X}$ we get a (1-α) confidence interval expressed as
$[\,\overline{X} - \Phi^{-1}(1-\alpha/2)\frac{\sqrt{S_{N-1}^2}}{\sqrt{N}}, \overline{X} + \Phi^{-1}(1-\alpha/2)\frac{\sqrt{S_{N-1}^2}}{\sqrt{N}}\,]$

Calculating in MatLab using (norminv) function we get [APPENDIX 9]

⇨ $\Phi^{-1}(1 - 0.05/2) = 1.96$

Now we calculate
$[\,\overline{X} - (1.96)\frac{\sqrt{S_{N-1}^2}}{\sqrt{N}}, \overline{X} + (1.96)\frac{\sqrt{S_{N-1}^2}}{\sqrt{N}}\,]$
⇨ [0.0167, 0.0132]

This is the 95% confidence interval of M

Now, we have $M = e^{-rt}$
⇨ $t = (-1/r)\ln(M)$

To calculate interval [t_1 , t_2] [APPENDIX 10]

$t_1 = (-1/r) . \ln (0.0167)$
⇨ $t_1 = 33837$ years

$t_2 = (-1/r) . \ln (0.0132)$
⇨ $t_2 = 35777$ years

Hence we get a 95% confidence interval of the time **[33837 , 35777]**

We find a time period dating from when the individual passed away to when these readings were taken. If we consider these readings to have been taken in 2010 we get

2010 – 33837 = - 31827 years
2010 – 35777 = - 33767 years

So we get that the individual died between 31827 BC – 33767 BC

Assuming that the average life span of a human in that era was 35 years
So,

He was born between 31792 BC – 33732 BC and died between 31827 BC – 33767 BC

ALTERNATIVELY (using error propagation)

We are given a data containing five measurements of M. Using that data, we first calculate the mean and standard deviation of the data.

We have N = Sample size = 5, and X_i = different values of M from the sample

In order to calculate the **mean** we use the expression
$\overline{X} = \frac{1}{N}\Sigma(X_i)$

⇨ $\overline{X} = 0.0149$ [APPENDIX 6]

Using an **unbiased estimate for the true variance** of X, the sample variance is given by
$\mathrm{Var}(X) \approx \sigma^2 = \frac{1}{N-1} [\left(\sum_{i=1}^{N} X_i \right) - N\overline{X}^2]$ [APPENDIX 7]

The **standard deviation** of the sample is given by
$SD(\overline{X}) = \sqrt{\frac{\sigma^2}{N}}$

⇨ $SD(\overline{X}) = 8.9028 \times 10^{-4}$ [APPENDIX 8]

Now using method of **error propagation**, we calculate Expectation, Variance and Standard Deviation of ‘t’

We have $M = e^{-rt}$

⇨ $t = (-1/r) \ln(M)$

[APPENDIX 18]

Expectation_t = 34750 years

Variance_t = 1213300 $years^2$

StandardDeviation_t = 1101.5 years

The 95% confidence interval calculated on Matlab gives us [APPENDIX 19]

t_1 = 33785 years

t_2 = 35716 years

Hence we get a 95% confidence interval of the time **[33785 , 35716]**

We find a time period dating from when the individual passed away to when these readings were taken. If we consider these readings to have been taken in 2010 we get

2010 – 33785 = -31775 years

2010 – 35716 = -33706 years

So we get that the individual died between 31775 BC – 33706 BC

Assuming that the average life span of a human in that era was 35 years

So,

He was born between 31740 BC – 33671 BC and died between 31775 BC – 33706 BC

MODEL 3:
PART 1.

(a) In order to find the likeliness of a flight being early rather than late we take an average of their probabilities of being early by 22 and 7 mins respectively.
The data calculated is shown in the table below

		% Early				
Destination City	**Airline**	**22 min**	**7 min**	**Mean Early %**		**Outcome**
BEIJING	AIR CHINA	14.03	32.04	46.07	44.93	Late
BEIJING	BRITISH AIRWAYS PLC	22.32	21.47	43.79		
BEIJING DAXING INT AIRPORT	BRITISH AIRWAYS PLC	9.3	21.71	31.01		Late
GUANGZHOU BAIYUN INT	CHINA SOUTHERN	24.03	43.49	67.52		Early
CHANGSHA HUANGHUA INT AIRPORT	HAINAN AIRLINES	12.59	39.12	51.71		Early
CHENGDU	AIR CHINA	11.74	31.3	43.04		Late
QINGDAO	BEIJING CAPITAL AIRLINES	30.08	42.37	72.45		Early
SHANGHAI (PU DONG)	BRITISH AIRWAYS PLC	9.12	31.91	41.03	41.49	Late
SHANGHAI (PU DONG)	CHINA EASTERN AIRLINES	5.8	30.39	36.19		
SHANGHAI (PU DONG)	VIRGIN ATLANTIC LTD	13.32	33.93	47.25		
SHENZHEN (HUANGTIAN)	SHENZHEN AIRLINES	7.57	26.5	34.07		Late
TIANJIN	TIANJIN AIRLINES	12.24	32.78	45.02		Late
WUHAN TIANHE INT	CHINA SOUTHERN	17.31	42.63	59.94		Early
ZHENGZHOU XINZHENG	CHINA SOUTHERN	4	36	40		Late

The airports for which the probability of being early is greater than 50% will have flights from Heathrow more likely to be early than late.

From the table we can observe that at the following airports flights arriving from Heathrow are more likely to be early than late

- ⇨ **GUANGZHOU BAIYUN INT**
- ⇨ **CHANGSHA HUANGHUA INT AIRPORT**
- ⇨ **QINGDAO**
- ⇨ **WUHAN TIANHE INT**

(b) The Expectation and the Standard Deviations were calculated for the three flights arriving at Shanghai Pu-Dong Airport as shown in the tables for each.

The percentage probabilities were divided by 100 to get the column for [f] and the time by which the flights arrived were taken to be negative whereas the time by which they were late were taken to be positive. The values of time together gave the column [x].

To calculate the expectation
First the column [fx] was calculated
The expectation [E] of a probability distribution is given by
E = Σ [fx]

To calculate the standard deviation
First a column corresponding to $[x-\mu]^2$ is calculated where μ is the mean (Expectation[E])
Then a column corresponding to $f.[x-\mu]^2$ is calculated
The variance σ^2 of a probability distribution is calculated using
$\sigma^2 = \Sigma f.[x-\mu]^2$
The standard deviation is given by σ ($\sqrt{\sigma^2}$)

Destination City	SHANGHAI (PU DONG)
Airline	BRITISH AIRWAYS PLC

Time Delayed (mins) [x]	Percentage Late/Early	[f]	[fx]	[x-μ]	[x-μ]^2	[f].[x-μ]^2
-22	9.12	0.0912	-2.0064	-34.3066	1176.9428	107.337184
-7	31.91	0.3191	-2.2337	-19.3066	372.744804	118.942867
8	36.97	0.3697	2.9576	-4.3066	18.5468036	6.85675328
23	10.8	0.108	2.484	10.6934	114.348804	12.3496708
46	6.54	0.0654	3.0084	33.6934	1135.2452	74.2450363
91	2.97	0.0297	2.7027	78.6934	6192.6512	183.921741
151	0.5	0.005	0.755	138.6934	19235.8592	96.179296
271	0.4	0.004	1.084	258.6934	66922.2752	267.689101
450	0.79	0.0079	3.555	437.6934	191575.512	1513.44655
		Σ [fx] =	12.3066		$\Sigma f.[x-\mu]^2$ =	2380.9682

Expectation [E]/μ = 12.3066 minutes

Standard Deviation = 48.79516571 minutes

Destination City SHANGHAI (PU DONG)
Airline CHINA EASTERN AIRLINES

Time Delayed (mins) [x]	Percentage Late/Early	[f]	[fx]	[x-μ]	[x- μ]^2	[f].[x- μ]^2
-22	5.8	0.058	-1.276	-37.7642	1426.134802	82.7158185
-7	30.39	0.3039	-2.1273	-22.7642	518.2088016	157.4836548
8	30.8	0.308	2.464	-7.7642	60.28280164	18.56710291
23	14.09	0.1409	3.2407	7.2358	52.35680164	7.377073351
46	12.02	0.1202	5.5292	30.2358	914.2036016	109.8872729
91	5.66	0.0566	5.1506	75.2358	5660.425602	320.3800891
151	0.69	0.0069	1.0419	135.2358	18288.7216	126.1921791
271	0.41	0.0041	1.1111	255.2358	65145.3136	267.0957858
450	0.14	0.0014	0.63	434.2358	188560.73	263.985022
		Σ [fx] =	15.7642		Σ f.[x-μ]^2 =	1353.683998

Expectation [E]/μ = 15.7642 minutes

Standard Deviation = 36.79244485 minutes

Destination City SHANGHAI (PU DONG)
Airline VIRGIN ATLANTIC LTD

Time Delayed (mins) [x]	Percentage Late/Early	[f]	[fx]	[x-μ]	[x- μ]^2	[f].[x- μ]^2
-22	13.32	0.1332	-2.9304	-30.041	902.461681	120.207896
-7	33.93	0.3393	-2.3751	-15.041	226.231681	76.7604094
8	33.52	0.3352	2.6816	-0.041	0.001681	0.00056347
23	8.65	0.0865	1.9895	14.959	223.771681	19.3562504
46	6.59	0.0659	3.0314	37.959	1440.88568	94.9543664
91	2.47	0.0247	2.2477	82.959	6882.19568	169.990233
151	0.96	0.0096	1.4496	142.959	20437.2757	196.197847
271	0.27	0.0027	0.7317	262.959	69147.4357	186.698076
450	0.27	0.0027	1.215	441.959	195327.758	527.384946
		Σ [fx] =	8.041		Σ f.[x-μ]^2 =	1391.55059

Expectation [E]/μ = 8.041 minutes

Standard Deviation = 37.30349297 minutes

Airline	Expectation [E]/µ (minutes)	Standard Deviation (minutes)
British Airways PLC	12.3066	48.795
China Eastern Airlines	15.7642	36.792
Virgin Atlantic LTD	8.041	37.303

According to the data calculated above, it is evident from the values of the expectations calculated that the Airline Virgin Atlantic is the most punctual with an expectation to be late by 8.041 minutes, followed by the Airline British Airways which is expected to be late by 12.3066 minutes. The airline China Eastern Airlines is expected to be the least punctual with and expectation to be late by 15.7642 minutes.

Ranking in order of punctuality

1. **Virgin Atlantic LTD**
2. **British Airways PLC**
3. **China Eastern Airlines**

However, these predictions, as is evident from the large values of standard deviation calculated from the data, are not very reliable.

For instance, the expectation of **British airways** to be late is by 12.3066 minutes but the standard deviation of this value is about 48.8 minutes which means that the airline can **vary between being late by 61.1066 minutes (48.8 + 12.3066) and being early by 36.4934 minutes (12.3066 - 48.8)**

Similarly, **China Eastern** can **vary between being late 52.5642 minutes (15.7642 + 36.8) and being early by 21.0358 minutes (15.7642 - 36.8)**.

And **Virgin Atlantic** can **vary between being late 45.341 minutes (8.041 + 37.3) and being early by 29.259 minutes (8.041 - 37.3).**

PART 2.

(a) Using passenger analysis data of the Heathrow airport from the internet, it was observed that the number of transit passengers comprised of 34% of the total number of passengers at the airport. [2]

Here we make a mathematical assumption that the overseas transit passengers comprise of 34% of the total overseas passengers and the home transit passengers comprise of 34% of the total home passengers.

FOR OVERSEAS
We use table 3 to calculate the total overseas passengers. This is done by taken the sum of the total passengers from Non-EU Europe, Africa, North America, Latin America, Middle East and Asia/Pacific for each month – which gives us the total overseas passengers for each month of the year 2019.

In order to get random variable $X_{3\text{-}os}$, we first multiply the total overseas passengers for each month by 0.34 (34%) and then divide it by the number of days in that month (31, 30 or 28) and then divide it by no. of hours in a day i.e. 24 to get the 'no. of overseas transit passengers per hour'

The mean of random variable $X_{3\text{-}os}$ is calculated using the formula
$\boldsymbol{\mu = (1/N)\, \Sigma\, X_{3\text{-}os}}$
where N=12 (sample size-12months)

We the subtract the mean from each element of random variable $X_{3\text{-}os}$ and then square it.
The variance is calculated using the formula
$\boldsymbol{Var(X) = \sigma^2 = (1/N)\, \Sigma\, (X_{3\text{-}os} - \mu)^2}$

[APPENDIX 11]

Month	Total Overseas	OS Transit Passengers (per hour) [$X_{3\text{-}os}$]	[$X_{3\text{-}os} - \mu$]	[$X_{3\text{-}os} - \mu$]^2
Jan	37,84,069	1729.278844	-157.592	24835.24
Feb	33,02,342	1670.827798	-216.0431	46674.6
Mar	39,18,856	1790.875054	-95.9958	9215.194
Apr	40,59,393	1916.935583	30.064729	903.8879
May	39,28,539	1795.300081	-91.57077	8385.207
Jun	42,78,749	2020.520361	133.64951	17862.19
Jul	46,17,842	2110.30414	223.43329	49922.43
Aug	46,23,740	2112.999462	226.12861	51134.15
Sept	40,54,959	1914.84175	27.970896	782.371
Oct	42,11,975	1924.827285	37.956431	1440.691
Nov	37,28,882	1760.860944	-126.0099	15878.5
Dec	41,46,441	1894.878952	8.0080976	64.12963
	$\Sigma X_{3\text{-}os}$ =	22642.45025	$\Sigma (X_{3\text{-}os} - \mu)^2$ =	227098.6

Mean (μ)	1886.870854 $hour^{-1}$
Variance (σ^2)	18924.88267 $hour^{-2}$

FOR HOME

We use table 3 to calculate the total home passengers. This is done by taken the sum of the total passengers from UK and EU for each month – which gives us the total home passengers for each month of the year 2019.

In order to get random variable X_{3-H} , we first multiply the total home passengers for each month by 0.34 (34%) and then divide it by the number of days in that month (31, 30 or 28) and then divide it by no. of hours in a day i.e. 24 to get the 'no. of home transit passengers per hour'

The mean of random variable X_{3-H} is calculated using the formula

$\mu = (1/N) \Sigma X_{3-H}$

where N=12 (sample size-12months)

We the subtract the mean from each element of random variable X_{3-H} and then square it.

The variance is calculated using the formula

$Var(X) = \sigma^2 = (1/N) \Sigma (X_{3-H} - \mu)^2$

[APPENDIX 12]

Month	Total Home	H Transit Passengers (per hour) [X_{3-H}]	[$X_{3-H} - \mu$]	[$X_{3-H} - \mu$]^2
Jan	21,43,474	979.5445699	-273.34891	74719.62502
Feb	21,79,946	1102.948869	-149.94461	22483.38545
Mar	26,08,633	1192.117231	-60.776246	3693.752056
Apr	27,38,819	1293.331194	40.437717	1635.208992
May	28,40,364	1298.015806	45.122329	2036.024615
Jun	29,67,848	1401.483778	148.5903	22079.07749
Jul	31,36,297	1433.254005	180.36053	32529.9202
Aug	30,56,587	1396.827392	143.93392	20716.97202
Sep	27,22,156	1285.462556	32.569079	1060.744878
Oct	28,52,804	1303.700753	50.807276	2581.379263
Nov	25,04,919	1182.878417	-70.01506	4902.108673
Dec	25,49,638	1165.157151	-87.736326	7697.662981
	ΣX_{3-H} =	15034.72172	$\Sigma (X_{3-H} - \mu)^2$ =	196135.8616

Mean (μ)	1252.893477 $hour^{-1}$
Variance (σ^2)	16344.65514 $hour^{-2}$

Final Result

Since a rate belonging in (persons/hour) cannot be in decimals (i.e., we cannot have a decimal value of a person) we round the values of mean and variance to the nearest whole number.

	Mean (μ) [$hour^{-1}$]	Variance (σ^2) [$hour^{-2}$]
$X_{3\text{-}os}$	1887 ($\mu_{X3\text{-}os}$)	18925 ($\sigma_{X3\text{-}os}^2$)
$X_{3\text{-}H}$	1253 ($\mu_{X3\text{-}H}$)	16345 ($\sigma_{X3\text{-}H}^2$)

(b) To determine the number of gates that need to be open in order to ensure the queues of passengers waiting to go through passport control does not increase with time.

$\frac{dN}{dt} = X_1 - X_2 - X_3$ ....(i)

For the queues of passengers waiting to go through passport control to not increase with time,

$$\frac{dN}{dt} \leq 0$$

However, in order to obtain a limiting value i.e. the maximum number of gates that need to stay open so that the queues of passengers does not increase with time we assume that,

$$\frac{dN}{dt} = 0$$

Plugging this value into equation (i) we have

$X_1 - X_2 - X_3 = 0$

⇨ $X_1 - X_3 = X_2$

Separating this equation for overseas and home passengers we have

⇨ $X_{1\text{-}os} - X_{3\text{-}os} = X_{2\text{-}os}$ and $X_{1\text{-}H} - X_{3\text{-}H} = X_{2\text{-}H}$ ....(ii)

Where $X_1 = X_{1\text{-}os} + X_{1\text{-}H}$, $X_2 = X_{2\text{-}os} + X_{2\text{-}H}$ and $X_3 = X_{3\text{-}os} + X_{3\text{-}H}$

We also have

$X_2 = n_{os}\, r_{os} + n_H\, r_H$

⇨ $X_{2\text{-}os} = n_{os}\, r_{os}$ and $X_{2\text{-}H} = n_H\, r_H$

Plugging these values into equations (ii)

For overseas

$X_{1\text{-}os} - X_{3\text{-}os} = n_{os}\, r_{os}$

⇨ $n_{os} = (X_{1\text{-}os} - X_{3\text{-}os}) / r_{os}$ ....(iii)

For home

$X_{1\text{-}H} - X_{3\text{-}H} = n_H\, r_H$

⇨ $n_H = (X_{1\text{-}H} - X_{3\text{-}H}) / r_H$ ....(iv)

In order to calculate the number of gates open for overseas and home passengers i.e., n_{os} and n_H we need to make assumptions for the values for r_{os} and r_H which is the number of people that pass through a till per hour.

From personal experience, it takes approximately 60 seconds for an overseas passenger and approximately 30 seconds for a home passenger to pass through a till.

⇨ $r_{os} = 3600/60 = 60$ hour^{-1} and $r_H = 3600/30 = 120$ hour^{-1}

For the values of $X_{3\text{-}os}$ and $X_{3\text{-}H}$ we use their means as calculated in (a)
And use mean values of$X_{1\text{-}os}$ [APPENDIX 13] and $X_{1\text{-}H}$ [APPENDIX 14]

As we use the means/expectations of $X_{1\text{-}os}$, $X_{3\text{-}os}$, $X_{1\text{-}H}$, and $X_{3\text{-}H}$ we apply the linear property of expectation which states that

$E[X_1 + X_2 + ... + X_n] = E[X_1] + E[X_2] + ... + E[X_3]$

$E[aX_1] = a\, E[X_1]$

Also, from the assumptions made in (a) we have $X_{3\text{-}os} = 0.34\, X_{1\text{-}os}$ and $X_{3\text{-}H} = 0.34\, X_{1\text{-}H}$

⇨ $\mu_{(X1\text{-}os\ -\ X3\text{-}os)} = \mu_{(0.66\ X1\text{-}os)} = (0.66)\, \mu_{X1\text{-}os} = 3633$ hour^{-1} [APPENDIX 16]

Similarly $\mu_{(X1\text{-}H\ -\ X3\text{-}H)} = (0.66)\, \mu_{X1\text{-}H} = 2432$ hour^{-1} [APPENDIX 16]

So, we can modify equations (iii) and (iv)

$n_{os} = (0.66)\, \mu_{X1\text{-}os} / r_{os}$ and $n_H = (0.66)\, \mu_{X1\text{-}H} / r_H$

Now we plug these values into equations we get [APPENDIX 15]

⇨ $n_{os} = 61.05$ gates and $n_H = 20.2667$ gates

Since the number of gates can't be a decimal value

⇨ **$n_{os} = 62$ gates and $n_H = 21$ gates**

Since we calculated the values of n using mean values, it is essential to calculate the variance in order to adapt to this variation to of people through time [Appendix 17]

$Var(n_{os}) = Var((0.66)\ \mu_{X1-os} / r_{os})$

⇨ $Var(n_{os}) = (0.66/r_{os})^2\ Var(\mu_{X1-os})$ [using $Var(aX) = a^2 Var(X)$]

⇨ $Var(n_{os}) = 19.8089$

$\sigma (n_{os}) = \sqrt{Var(n_{os})} = 4.45$

⇨ **$\sigma (n_{os}) = 4.45 \sim 5$** (gates can't have decimal value)

$Var(n_H) = Var((0.66)\ \mu_{X1-H} / r_H)$

⇨ $Var(n_H) = (0.66/r_H)^2\ Var(\mu_{X1-H})$

⇨ $Var(n_H) = 4.2770$

$\sigma (n_H) = \sqrt{Var(n_H)} = 2.068$

⇨ **$\sigma (n_H) = 2.068 \sim 3$** (gates can't have decimal value)

Hence if we include the standard deviations in the values calculated for n_{os} and n_H

⇨ **$n_{os} = 62 \pm 5$ gates and $n_H = 21 \pm 3$ gates**

(c) We are required to predict how passport control queues change as a function of the number and type of tills open

Plugging in the value of $X_2 = n_{os}\ r_{os} + n_H\ r_H$ into equation (i) we get

$$\frac{dN}{dt} = X_1 - (n_{os}\ r_{os}\ +\ n_H\ r_H) - X_3$$

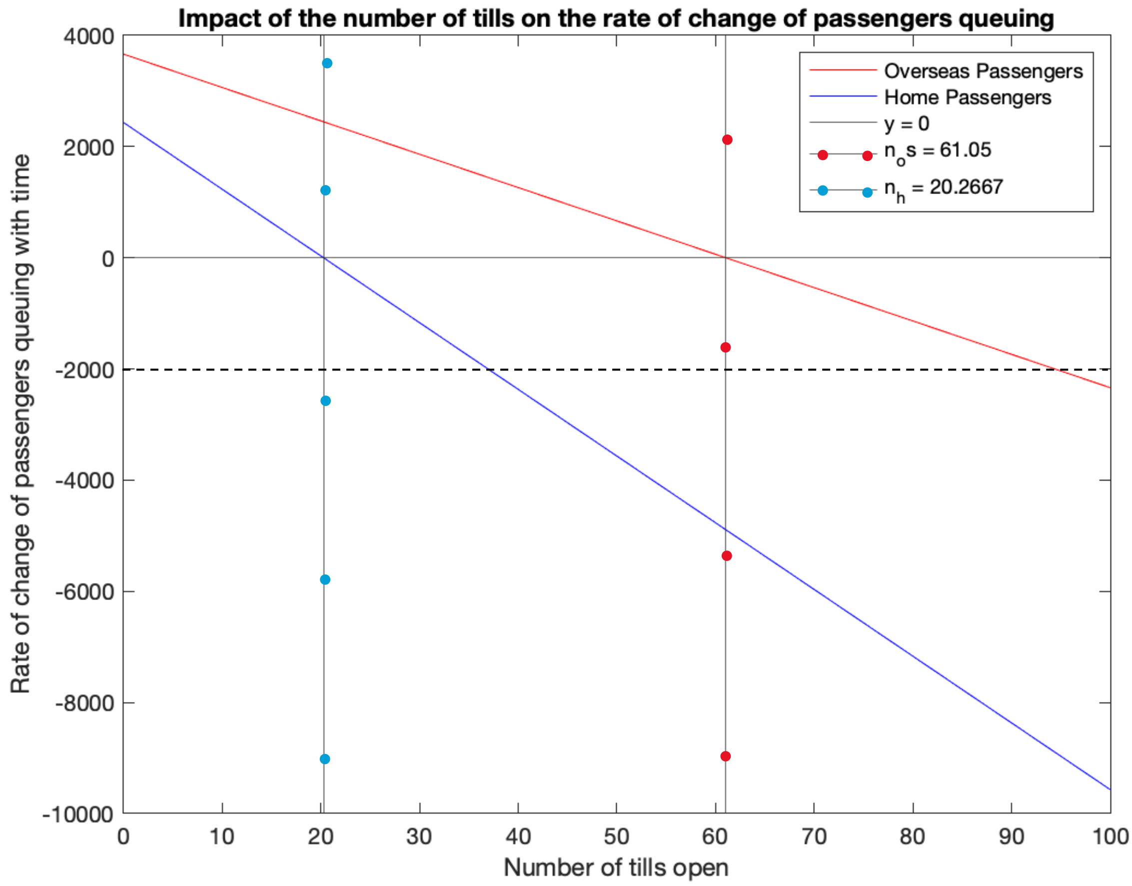
⇨ $$\frac{dN}{dt} = X_1 - n_{os}\ r_{os} - \ n_H\ r_H - X_3$$

To determine how the rate of change of passengers queuing with respect to time changes as a function of the number of rows we separate this equation into distinct equations for overseas and home passengers.

⇨ $$\frac{\mathbf{dN_{os}}}{\mathbf{dt}} = \mathbf{X_{1-os} - n_{os}\ r_{os} - X_{3-os}} \quad \text{and} \quad \frac{\mathbf{dN_H}}{\mathbf{dt}} = \mathbf{X_{1-H} - n_H\ r_H - X_{3-H}}$$

We plot these two equations with respect to the number of tills open

In order to get a broad range for possible value of the number of tills we chose it to be an array of [0:100] [APPENDIX 20]



After plotting the graph a number of observations were made

At dN/dt = 0

We start by observing where the lines of the graphs meet the x axis. These are the points where the Rate of change of passengers queuing with time is 0, i.e., the number of passengers queuing is constant. In such a condition the number of tills open for overseas is 61.05~62 and the number of tills open for home is 20.2667~21.

At dN/dt > 0 – queues get longer

This region in the graph is represented by the region above the line y =0. We can observe that as the value of dN/dt approaches 0, the no. of gates increases linearly. It can also be observed that no. of gates open increase more for overseas passengers than for home passengers as is evident from the steeper slope of home passengers in comparison to the overseas passengers.

At dN/dt < 0 – queues get shorter

This region in the graph is represented by the region below the line y =0. We can observe that as the value of dN/dt decreases from 0, the no. of gates increases linearly. It can also be observed that no. of gates open increase more for overseas passengers than for home passengers. This is evident from the fact that as dN/dt decreases from 0 to -2000 (dotted line on graph), the no. of gates for overseas passengers increases from about 60 to 95 (increase by 35 gates) whereas for home passengers it increases from about 20 to 35 (increase by 15 gates).

UNCERTAINITY OF MODEL

Since here we are considering random variables X_1, X_2 and X_3 in this model there is a level of uncertainty, it is important to consider the error propagation in the calculations. These errors must be propagated onto the value of the number of gates to opened calculated. This helps in estimating probable 'worst case scenarios' so that suitable adjustments can be made.

Furthermore, the model is based on educated assumptions made in order to estimate the values of certain constants (eg, r_{os} and r_H). The accuracy of these values is questionable which makes the model uncertain. Also these values are not constants in the real world.

Also, we have only studied the data of passengers at the airport for one year and there is no set formula that helps us predict the number of passengers at the airport at any given day making these values absolutely random. Hence to come up with a completely certain model of such unpredictable variables is improbable.

From the data that the total transit passengers is 34% of the total passengers, we made an assumption that the total overseas and home transit passengers would also be 34% of the total overseas and home passengers respectively. Since this was a personal assumption and taken from concrete data it make also contribute towards the uncertainty of the model.

AIRPORT STRATEGY

The airport needs to opt for a strategy to optimize the number of gates open in accordance with the number of passengers in the queue. The number of gates open shouldn't be too less as it may lead to a pile up of people in the queues and if the number of gates open are more than necessary it would lead to a waste of the resources of the airport.

It is evident from the graphs that the number of people queuing is directly proportional to the number of gates.
We have assumed $r_{os} = 60\ hr^{-1}$ and $r_H = 120\ hr^{-1}$
We can make an approximate relationship between N and n using r such that

$N_{os} = 60\ n_{os}$ and $N_H = 120\ n_H$

According to this relationship if the number of people queueing increases by say ΔN , the number of tills should increase by (1/60) ΔN for overseas and (1/120) ΔN for home.

APPENDIX:

1. Start by defining the values of variables

Incoming flow rate of water (r_i)

Leaving flow rate of water (r_o)

Concentration of incoming pollutants (c_i)

Volume of water in the lake

```
V = 13.08*10^6
```

```
V = 13080000
```

```
r_i = 45000
```

```
r_i = 45000
```

```
r_o = r_i
```

```
r_o = 45000
```

```
c_i = 450
```

```
c_i = 450
```

```
t= 0:2000
```

```
t = 1×2001
      0      1      2      3      4      5      6      7      8      9     10     11
12 ⋯
```

Now, plotting concentration of pollutants against time in order to determine the time (t) at which the damage becomes irreversible

```
c_o = (-(c_i)*V*exp(-r_o*t/V) + c_i*V)/V
```

```
c_o = 1×2001
          0     1.5455     3.0857     4.6206     6.1502     7.6746     9.1938 ⋯
```

```
plot(c_o,"R")
hold on
```

We now need to find the time when the concentration reached 331 g/m^3

Let this be the critical concentration (c_x) and the time corresponding be t_x

```
c_x =331
```

```
c_x = 331
```

```
t_x = (-V/r_o)*(log(1-(c_x/c_i)))
```

```
t_x = 386.6227
```

2. plotting the graph

```
plot(t_x, c_x, "ro")
xlabel ('time (hours)')
ylabel ('concentration of pollutants in the lake (g/m^3)')
hold on
```

3. We need to find the concentration that is 95% of the steady state concentration

Since this the concentration at time=tau is can be called c_tau

```
c_tau = 0.95*c_i
```

```
c_tau = 427.5000
```

We found the equation for tau

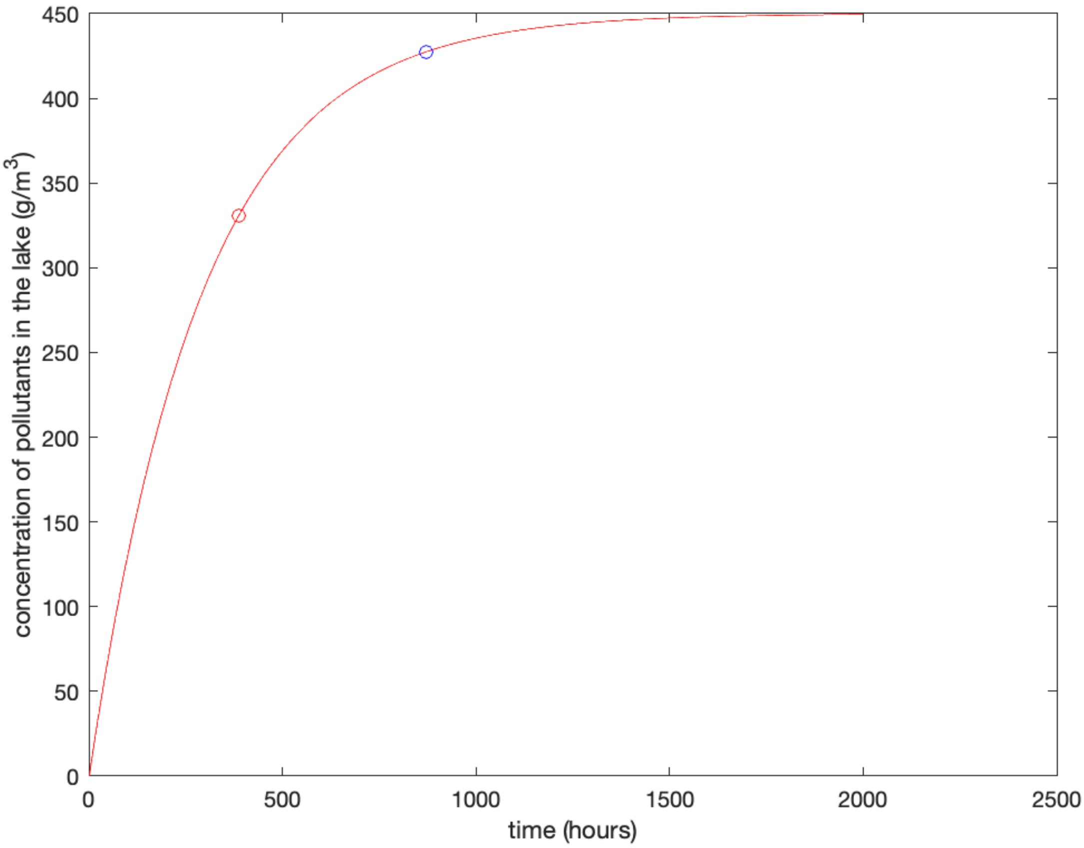
```
t = 0:2000
```

```
t = 1×2001
     0     1     2     3     4     5     6     7     8     9    10    11
12 ⋯
```

```
tau = (-V/r_o)*(log(1-(c_tau/c_i)))
```

```
tau = 870.7595
```

```
plot (tau, c_tau , "bo")
hold off
```



4. New initial concentration of the lake

```
c_L = 213
```

```
c_L = 213
```

Now, defining concentration of pollutant in the lake when there is now inflow of pollutants in the lake

```
c_o_new = c_L*exp((-r_o/V)*t)
```

```
c_o_new = 1×2001
  213.0000  212.2685  211.5394  210.8129  210.0889  209.3673  208.6483 ···
```

```
plot (t,c_o_new,'green')
hold on
```

Now we define the final desired concentration of the lake of 5 g/m^3

```
c_final = 5
```

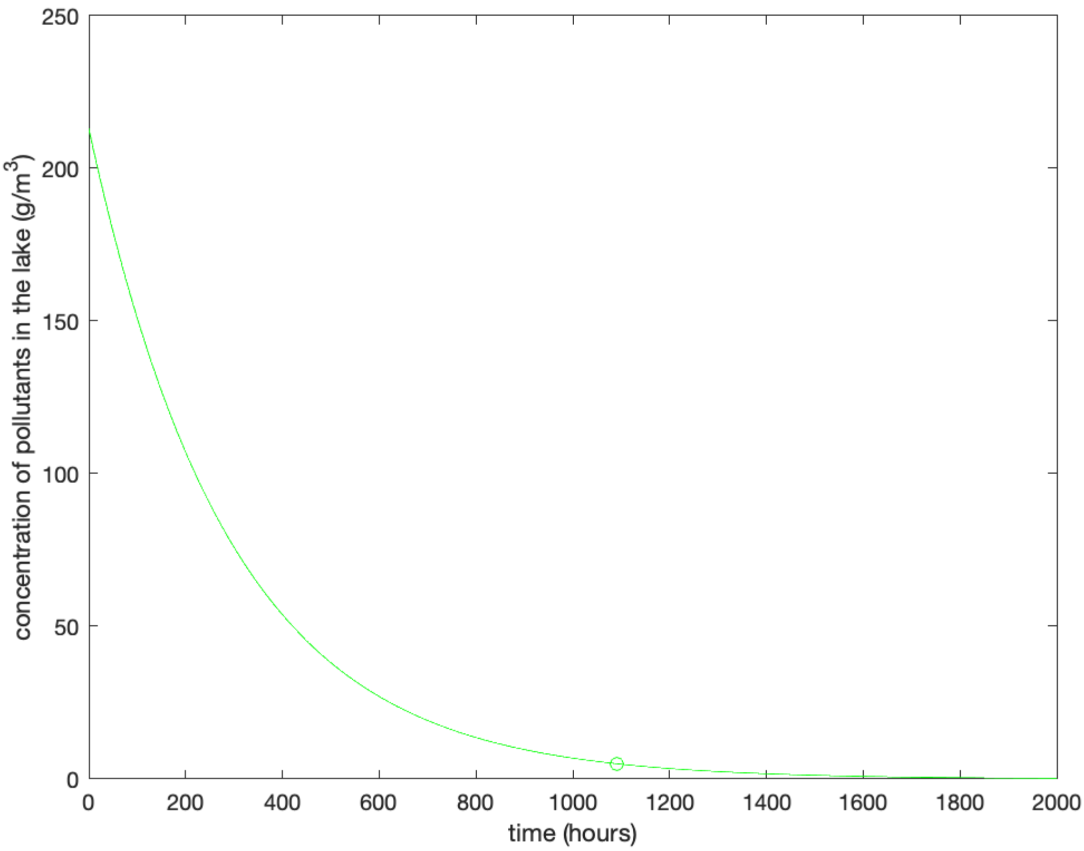
```
c_final = 5
```

Now finding the time to eliminate pollutants from the lake

```
t_eliminate = (-V/r_o)*log(c_final/c_L)
```

```
t_eliminate = 1.0905e+03
```

```
plot (t_eliminate , c_final , "go")
xlabel ('time (hours)')
ylabel ('concentration of pollutants in the lake (g/m^3)')
hold off
```



5. Checking whether the value of y(t)/V is less that 5 for t_eliminate

```
Finalconcoflakei = (c_L)*exp(-r_o*(t_eliminate)/V)
```

```
Finalconcoflakei = 5.0000
```

So in order to get final conc. less than 5 we increase t_eliminate by a small amount

```
Finalconcoflakeii = (c_L)*exp(-r_o*(1090.55)/V)
```

```
Finalconcoflakeii = 4.9998
```

6. We start by defining the different values of M from the sample provided in Table 1

```
M_1 = 0.0120
```

```
M_1 = 0.0120
```

```
M_2 = 0.0140
```

```
M_2 = 0.0140
```

```
M_3 = 0.0167
```

```
M_3 = 0.0167
```

```
M_4 = 0.0167
```

```
M_4 = 0.0167
```

```
M_5 = 0.0153
```

```
M_5 = 0.0153
```

```
N = 5
```

```
N = 5
```

```
r = 1.2097*10^(-4)
```

```
r = 1.2097e-04
```

Now we find the Mean of the values defined above

```
Mean = (M_1+M_2+M_3+M_4+M_5)/N
```

```
Mean = 0.0149
```

7. Now we calculate the unbiased variance of the data

```
Variance = (1/(N-1))*(((M_1)^2+(M_2)^2+(M_3)^2+(M_4)^2+(M_5)^2)-(N*Mean^2))
```

```
Variance = 3.9630e-06
```

8. We now calculate the standard deviation

```
StandardDeviation = sqrt(Variance/N)
```

```
StandardDeviation = 8.9028e-04
```

9. We now calculate the 95% confidence interval

```
n_inv = norminv (0.975,0,1)
```

```
n_inv = 1.9600
```

```
X_dash_1 = Mean + (((n_inv)*(StandardDeviation)))
```

```
X_dash_1 = 0.0167
```

```
X_dash_2 = Mean - (((n_inv)*(StandardDeviation)))
```

```
X_dash_2 = 0.0132
```

10. Now we convert this interval in M using the equation t = (-1/r)*log(M) to an interval in t

```
t_dash_1 = (-1/r)*log(X_dash_1)
```

```
t_dash_1 = 3.3837e+04
```

```
t_dash_2 = (-1/r)*log(X_dash_2)
```

```
t_dash_2 = 3.5777e+04
```

Thus the interval is 33837< t < 35777

11.

Month	Total Overseas	Transit Passengers (per month)	Transit Passengers (per day)	Transit Passengers (per hour) [x]	[x-u]	[x-u]^2
Jan	37,84,069	1286583.46	41502.69226	1729.278844	-157.5920099	24835.242
Feb	33,02,342	1122796.28	40099.86714	1670.827798	-216.0430564	46674.602
Mar	39,18,856	1332411.04	42981.00129	1790.875054	-95.99580024	9215.1937
Apr	40,59,393	1380193.62	46006.454	1916.935583	30.06472933	903.88795
May	39,28,539	1335703.26	43087.20194	1795.300081	-91.57077335	8385.2065
Jun	42,78,749	1454774.66	48492.48867	2020.520361	133.6495071	17862.191
Jul	46,17,842	1570066.28	50647.29935	2110.30414	223.4332858	49922.433
Aug	46,23,740	1572071.6	50711.9871	2112.999462	226.1286084	51134.148
Sept	40,54,959	1378686.06	45956.202	1914.84175	27.970896	782.37102
Oct	42,11,975	1432071.5	46195.85484	1924.827285	37.95643095	1440.6907
Nov	37,28,882	1267819.88	42260.66267	1760.860944	-126.0099096	15878.497
Dec	41,46,441	1409789.94	45477.09484	1894.878952	8.008097613	64.129627
				22642.45025		227098.59

Mean (u)	1886.870854
Variance (sigma)	18924.88267
Standard Deviati	137.5677385

12.

Month	Total Home	Transit Passengers (per month)	Transit Passengers (per day)	Transit Passengers (per hour) [x]	[x-u]	[x-u]^2
Jan	21,43,474	728781.16	23509.06968	979.5445699	-273.34891	74719.62502
Feb	21,79,946	741181.64	26470.77286	1102.948869	-149.94461	22483.38545
Mar	26,08,633	886935.22	28610.81355	1192.117231	-60.776246	3693.752056
Apr	27,38,819	931198.46	31039.94867	1293.331194	40.437717	1635.208992
May	28,40,364	965723.76	31152.37935	1298.015806	45.122329	2036.024615
Jun	29,67,848	1009068.32	33635.61067	1401.483778	148.5903	22079.07749
Jul	31,36,297	1066340.98	34398.09613	1433.254005	180.36053	32529.9202
Aug	30,56,587	1039239.58	33523.85742	1396.827392	143.93392	20716.97202
Sep	27,22,156	925533.04	30851.10133	1285.462556	32.569079	1060.744878
Oct	28,52,804	969953.36	31288.81806	1303.700753	50.807276	2581.379263
Nov	25,04,919	851672.46	28389.082	1182.878417	-70.01506	4902.108673
Dec	25,49,638	866876.92	27963.77161	1165.157151	-87.736326	7697.662981
				15034.72172		196135.8616

Mean (u)	1252.893477
Variance (sigma)	16344.65514
Standard Deviati	127.8462167

13. Excel Table to find mean and variance of $X_{1\text{-}os}$

Month	Total Overseas	Total Overseas per day	Total Overseas per hour (x)	x-u	(x-u)^2
Jan	37,84,069	122066.7419	5086.114247	-463.5059127	214837.7311
Feb	33,02,342	117940.7857	4914.199405	-635.4207552	403759.5362
Mar	39,18,856	126414.7097	5267.27957	-282.3405901	79716.20882
Apr	40,59,393	135313.1	5638.045833	88.42567333	7819.099704
May	39,28,539	126727.0645	5280.294355	-269.3258052	72536.38933
Jun	42,78,749	142624.9667	5942.706944	393.0867844	154517.2201
Jul	46,17,842	148962.6452	6206.776882	657.1567217	431854.9569
Aug	46,23,740	149152.9032	6214.704301	665.0841411	442336.9147
Sept	40,54,959	135165.3	5631.8875	82.26734	6767.915231
Oct	4211975	135870.1613	5661.25672	111.6365604	12462.72162
Nov	3728882	124296.0667	5179.002778	-370.6173822	137357.244
Dec	4146441	133756.1613	5573.173387	23.5532271	554.7545067
			66595.44192		1964520.692
MEAN (u)	5549.62016				
Var(uX1os)	163710.0577				
sd	404.6109955				

$\mu_{X1\text{-}os}$ ← MEAN (u) = $(1/N)\Sigma x$

Var(uX1os) = $(1/N)\Sigma (x-u)^2$ → $Var(\mu_{X1\text{-}os})$

14. Excel Table to find mean and variance of $X_{1\text{-}H}$

Month	Total Home	Total Home per day	Total Home per hour (x)	x-u	(x-u)^2
Jan	2143474	69144.32258	2881.013441	-803.9674691	646363.6914
Feb	2179946	77855.21429	3243.967262	-441.0136481	194493.0378
Mar	2608633	84149.45161	3506.227151	-178.7537595	31952.90652
Apr	2738819	91293.96667	3803.915278	118.9343678	14145.38384
May	2840364	91624.64516	3817.693548	132.7126384	17612.64439
Jun	2967848	98928.26667	4122.011111	437.0302011	190995.3967
Jul	3136297	101170.871	4215.452957	530.472047	281400.5926
Aug	3056587	98599.58065	4108.31586	423.3349502	179212.4801
Sept	2722156	90738.53333	3780.772222	95.79131222	9175.975497
Oct	2852804	92025.93548	3834.413978	149.4330685	22330.24196
Nov	2504919	83497.3	3479.054167	-205.9267433	42405.82362
Dec	2549638	82246.3871	3426.932796	-258.0481143	66588.82929
			44219.76977		1696677.004
MEAN (u)	3684.980814				
var(uX1H)	141389.7503				
sd	376.0182845				

$\mu_{X1\text{-}H}$ ← MEAN (u) = $(1/N)\Sigma x$

var(uX1H) = $(1/N)\Sigma (x-u)^2$ → $Var(\mu_{X1\text{-}H})$

15. To determine the number of gates that need to be open in order to ensure the queues of passengers waiting to go through passport control does not increase with time

We start by defining the values for the total passengers (overseas)

```
Xos_1 = 5550
```

```
Xos_1 = 5550
```

the total passengers (home)

```
Xh_1 = 3685
```

```
Xh_1 = 3685
```

the transit passengers (overseas)

```
Xos_3 = 1887
```

```
Xos_3 = 1887
```

the transit passengers (home)

```
Xh_3 = 1253
```

```
Xh_3 = 1253
```

Now, defining rate at which overseas passengers pass through the gates per hour (r_os)

Assuming each passenger takes 1 minute (60seconds)

```
r_os = 60
```

```
r_os = 60
```

Defining rate at which home passengers pass through the gates per hour (r_h)

Assuming each passenger takes 30 seconds

```
r_h = 120
```

```
r_h = 120
```

we have n_os as the no. of gates open for overseas passengers and n_h as the number of gates open for home passengers

```
n_os = (Xos_1 - Xos_3)/r_os
```

```
n_os = 61.0500
```

```
n_h = (Xh_1 -Xh_3)/r_h
```

```
n_h = 20.2667
```

16. Calculations for the values of

```
xos = 0.66*Xos_1
```

```
xos = 3663
```

```
xh = 0.66*Xh_1
```

```
xh = 2.4321e+03
```

17. Now we calculate the variance of the no. of gates open

```
Var_uX1os = 163710.0577
```

Var_uX1os = 1.6371e+05

```
Var_nos = (0.66/r_os)^2*(Var_uX1os)
```

Var_nos = 19.8089

```
Var_uX1H = 141389.7503
```

Var_uX1H = 1.4139e+05

```
Var_nh = (0.66/r_h)^2*(Var_uX1H)
```

Var_nh = 4.2770

18. We first determine the mean time

```
E_t = -log(Mean)/r
```

E_t = 3.4750e+04

We then determine the variance of time

```
Var_t = Variance*(-1/(Mean*r))^2
```

Var_t = 1.2133e+06

We then determine the standard deviation

```
SD_t = sqrt(Var_t)
```

SD_t = 1.1015e+03

19. Now we calculate the 95% confidence interval

```
a = 1-0.95
```

a = 0.0500

```
t_1 = norminv((a/2),E_t, SD_t/sqrt(N))
```

t_1 = 3.3785e+04

```
t_2 = norminv((1-a/2),E_t, SD_t/sqrt(N))
```

t_2 = 3.5716e+04

20. We want to predict how passport control queues change as a function of the number and types of tills open

The rate at which the passengers pass through the gate is defined for overseas and home as r_os and r_h respectively

The rate of change of passengers queuing with respect to time is defined as the following for

Overseas

```
dNosdt = Xos_1 - (r_os*n_os) - Xos_3
```

```
dNosdt = 1×2
        3663        -2337
```

Home

```
dNhdt = Xh_1 - (r_h*n_h) - Xh_3
```

```
dNhdt = 1×2
        2432        -9568
```

We range the number of tills open as an array as follows in order to plot the rate of change of passengers queuing with respect to time as a function of number of tills open

```
n_os = [0 100]
```

```
n_os = 1×2
     0    100
```

```
n_h = [0 100]
```

```
n_h = 1×2
     0    100
```

In order to compare we plot both the functions on the same graph

```
plot (n_os,dNosdt,'r')
hold on
plot (n_h,dNhdt,'b')
xlim([0.0 100.0])
ylim([-10000 4000])
title ('Impact of the number of tills on the rate of change of passengers queuing')
xlabel ('Number of tills open')
ylabel ('Rate of change of passengers queuing with time')

legend('Overseas Passengers','Home Passengers')
```

To show the points at which the slopes of the graphs meet the x axis we plot a line y=0

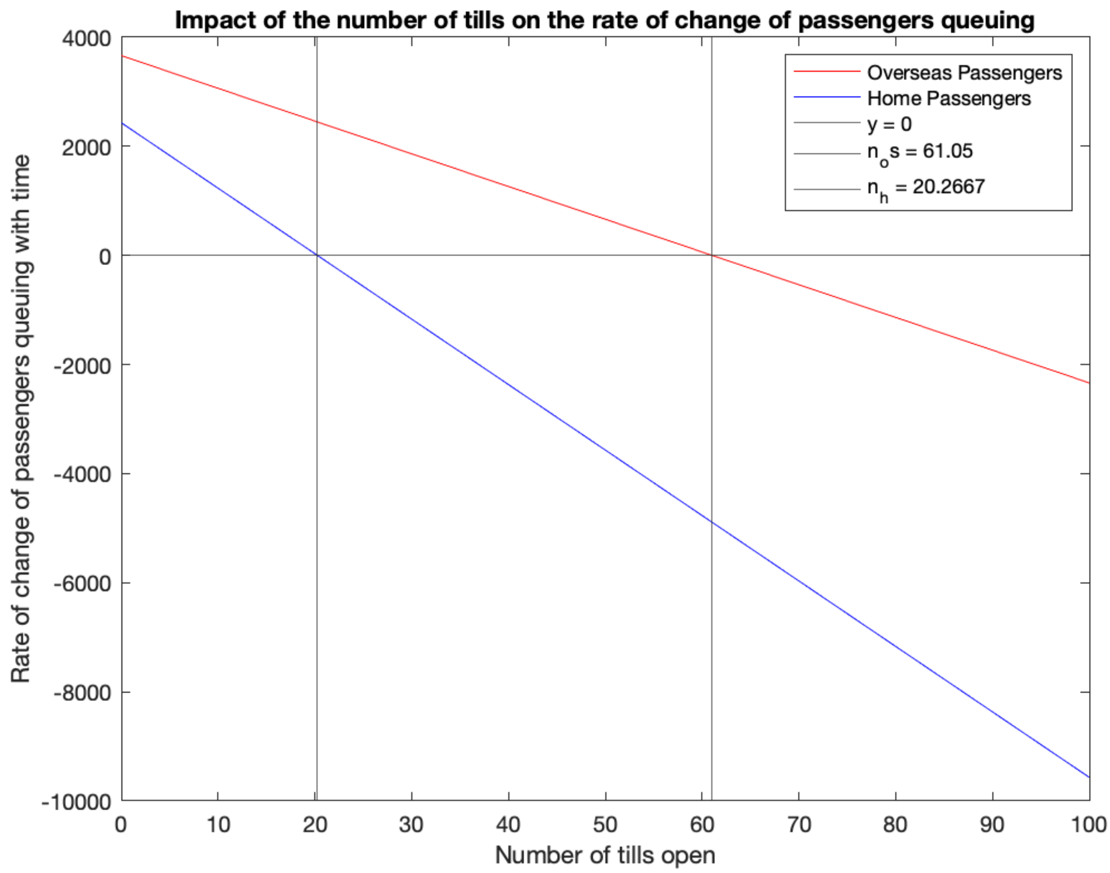
```
yline(0)
```

We know from the assumption made in part (b) that at the values of n calculated in part (b) the value of dN/dt =0

Hence, in order to mark the regions at which dN/dt are positive and negative we plot lines parallel to the y axis and passing through the values of n_os and n_h found in part (b)

```
xline (61.05)
xline (20.2667)

xlim([0.0 100.0])
ylim([-10000 4000])
```



REFERENCES:

1. Wikipedia. (2020). *Lake Pichola*. [online] Available at: https://en.wikipedia.org/wiki/Lake_Pichola [Accessed 5 Feb. 2021].

2. Statista. (n.d.). *UK: flight transfers at Heathrow Airport 2002-2019*. [online] Available at: https://www.statista.com/statistics/303920/flight-transfers-passengers-terminating-at-heathrow-airport-uk/.