Walid JAOUI – Youssef MARZOUK

Report

Practical work 1 : Parameter estimation

-

Marc FISCHER

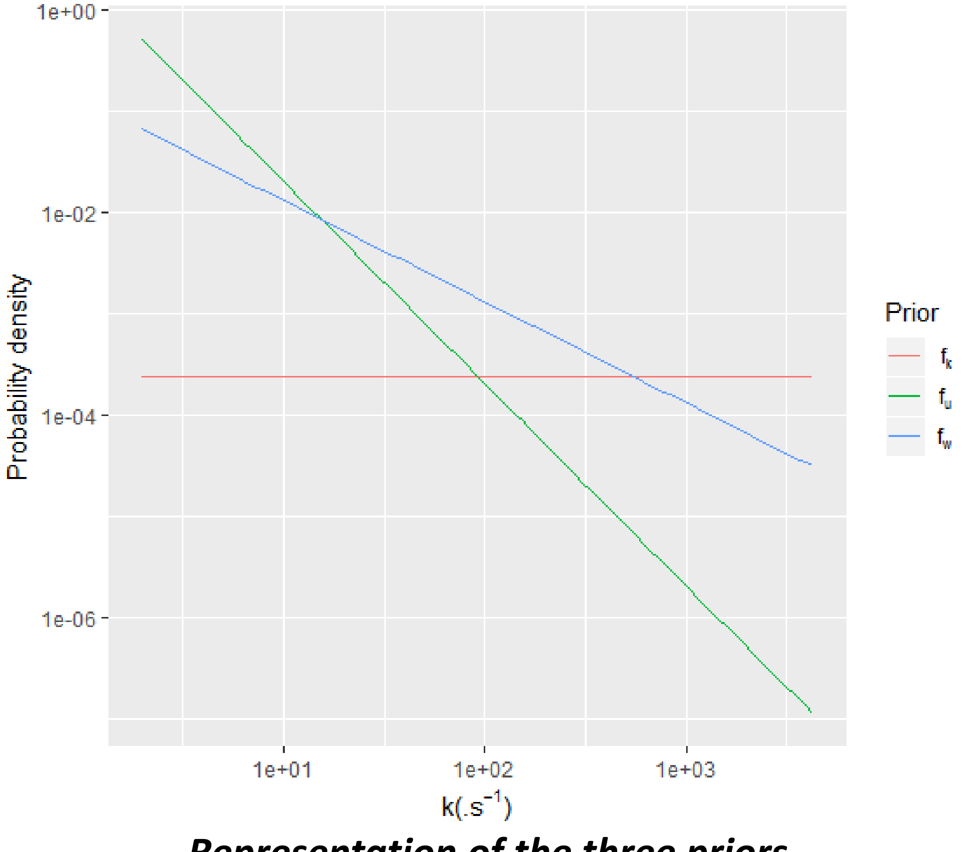
2019

I. First Model

Question 2 :

By plotting the three priors we notice that the 3 values of the returned vector are almost equal to 1.

```
> test_produce_priors_M1(2,4.26E+3,100)
[1] "Generate grid - n_k : " "100"
[1] "3 integrals : "    "0.999999999999999" "1.0029978753471"   "1.00099909220798"
[1] 1.000000 1.002998 1.000999
```



Representation of the three priors

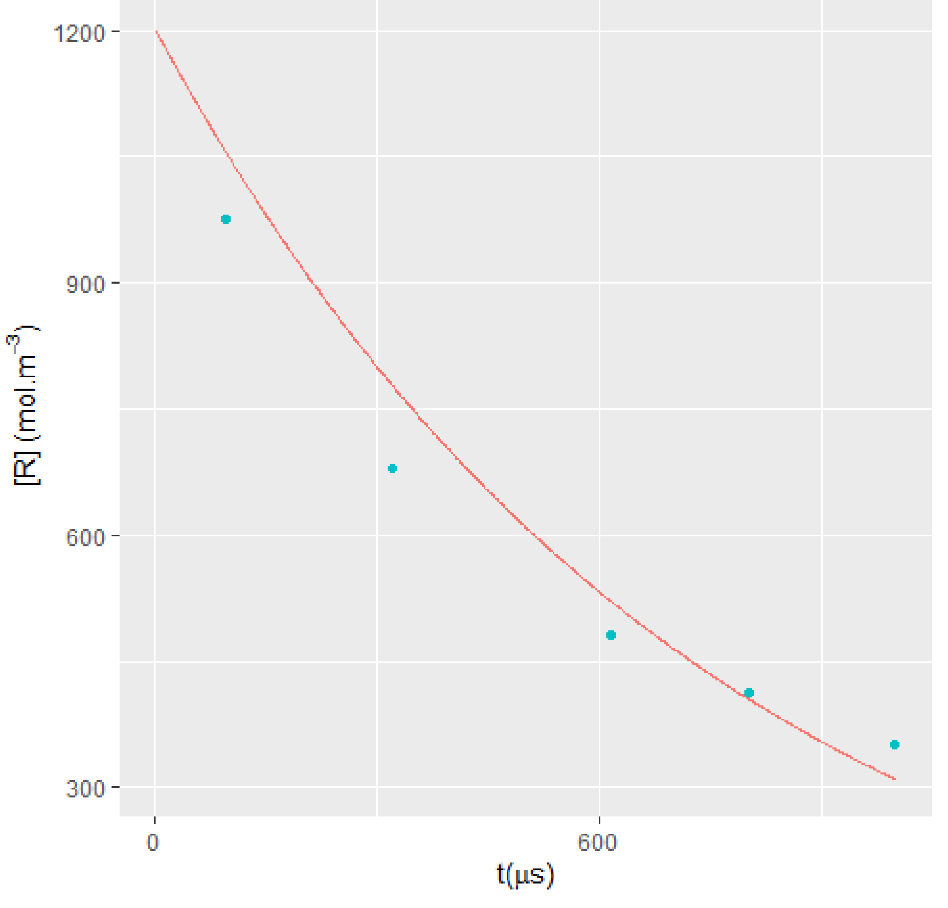
I. First Model

Question 3 :

In order to draw the profil for the first model, we use the following code:

```
R0= 1200
k=1.357e+03
t_mod=  seq(1,1000,1)
R_mod= Compute_R_profile_M1(t_mod,R0,k)
plot_profile_M1(t_mod,R_mod)
```

As a result, we obtain:



We notice that the curve does not go through all the experimental points

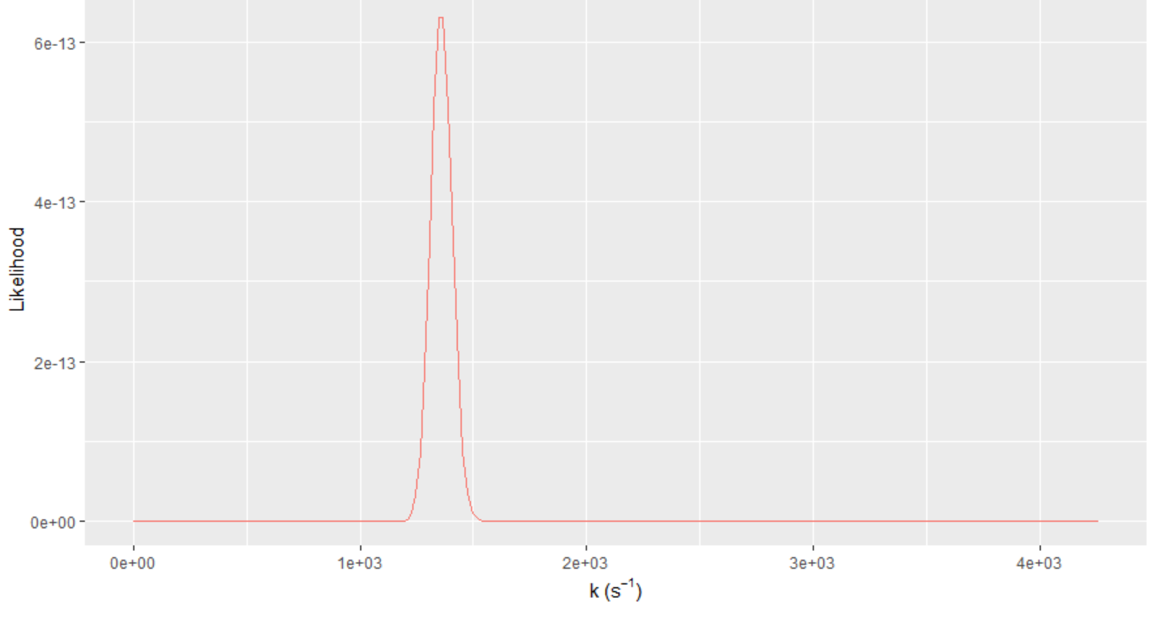
I. First Model

Question 4 :

Using the following code :

```
R0= 1200
lbound=1.2E+3
ubound=1.55E+3
n_k=500
Examine_likelihood_M1(R0,lbound,ubound,n_k)
```

We obtain the plot for the likelihood function:



We notice, according to the result that the optimal point is approximately k = **1.357e+03** /s.

```
[1] "Max lilelihood: "              "1356.49324348787"
[3] "0.000000000000645778428456967"
```

Result of the maximum likelihood according to a program coded in R

I. First Model

Question 4 :

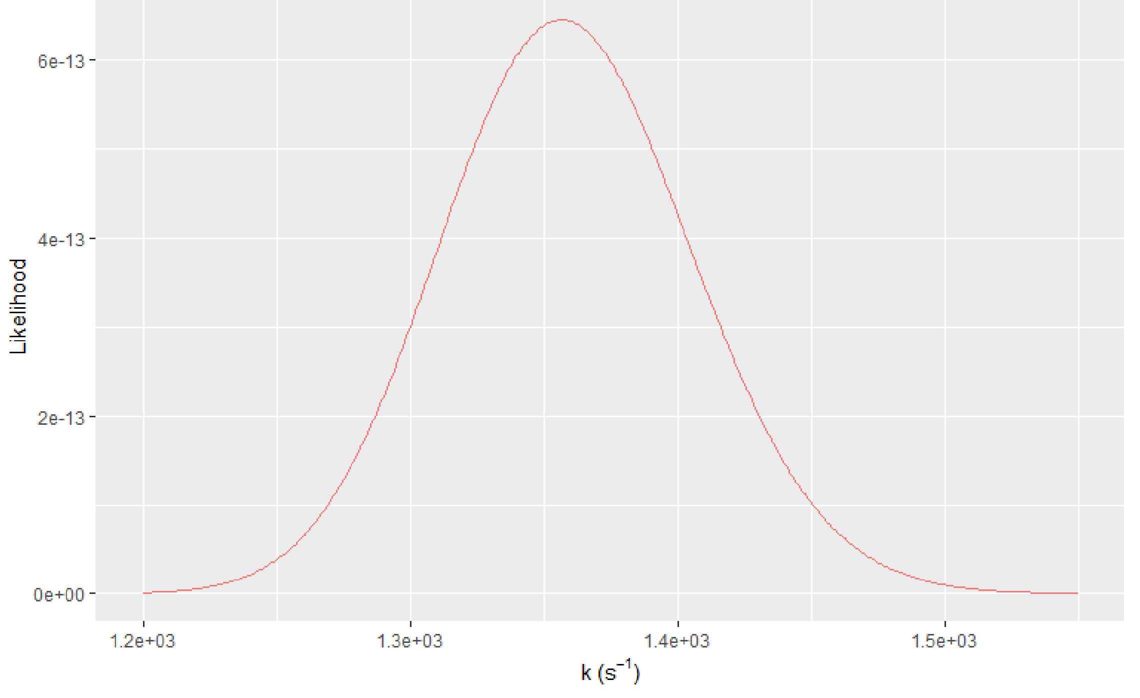
So, according to the optimal point, we narrow the interval and we choose a specific one where the likelihood has the strongest values for producing the graph.

For this, we choose the interval [1.2 E+3;1.55 E+3].

Using the following code:

```
R0= 1200
k=1.357e+03
lbound=1.2E+3
ubound=1.55E+3
n_k=500
Examine_likelihood_M1(R0,lbound,ubound,n_k)
```

We obtain the following plot:



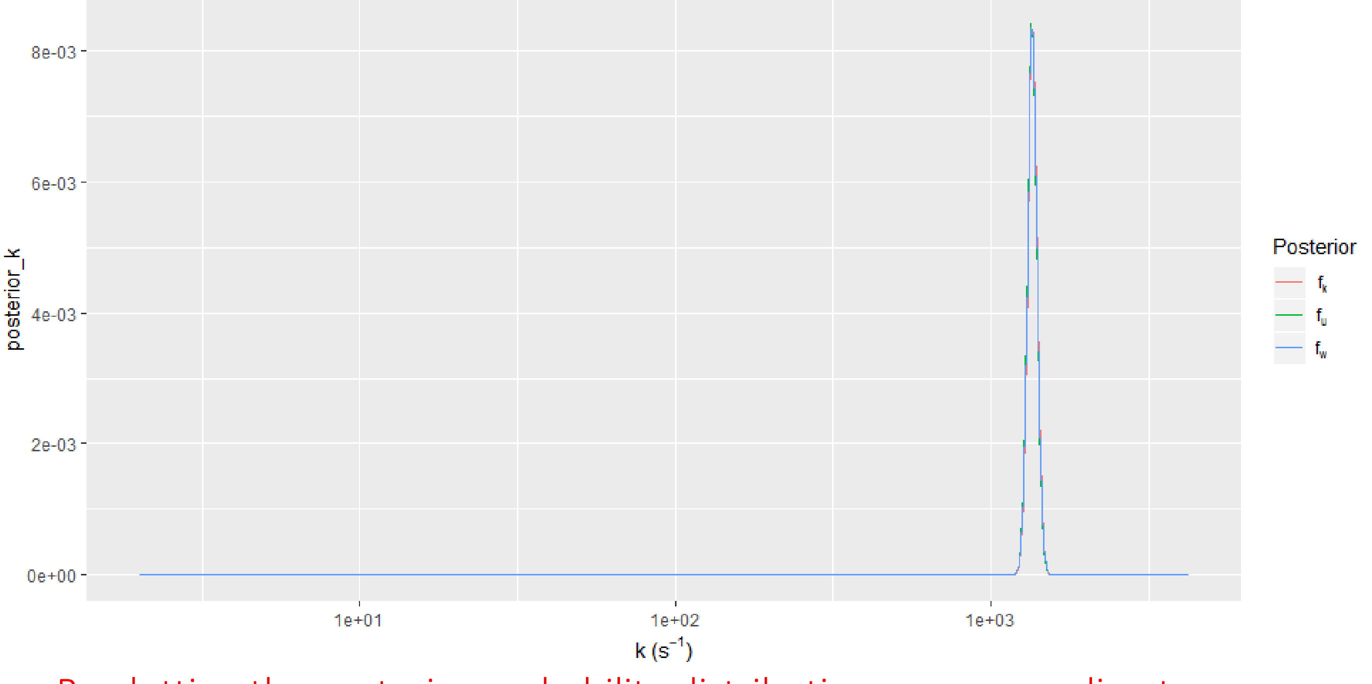
I. First Model

Question 5 :

To obtain the posterior probability distributions corresponding to the three priors, we use the following code :

```
R0= 1200
lbound=2
ubound=4.26E+3
eps = 0.065
lb_plot=lbound
ub_plot=ubound
n_k=500
Compute_all_posteriors_M1(R0,lbound,ubound,n_k,eps,lb_plot,ub_plot)
```

So we obtain the result:



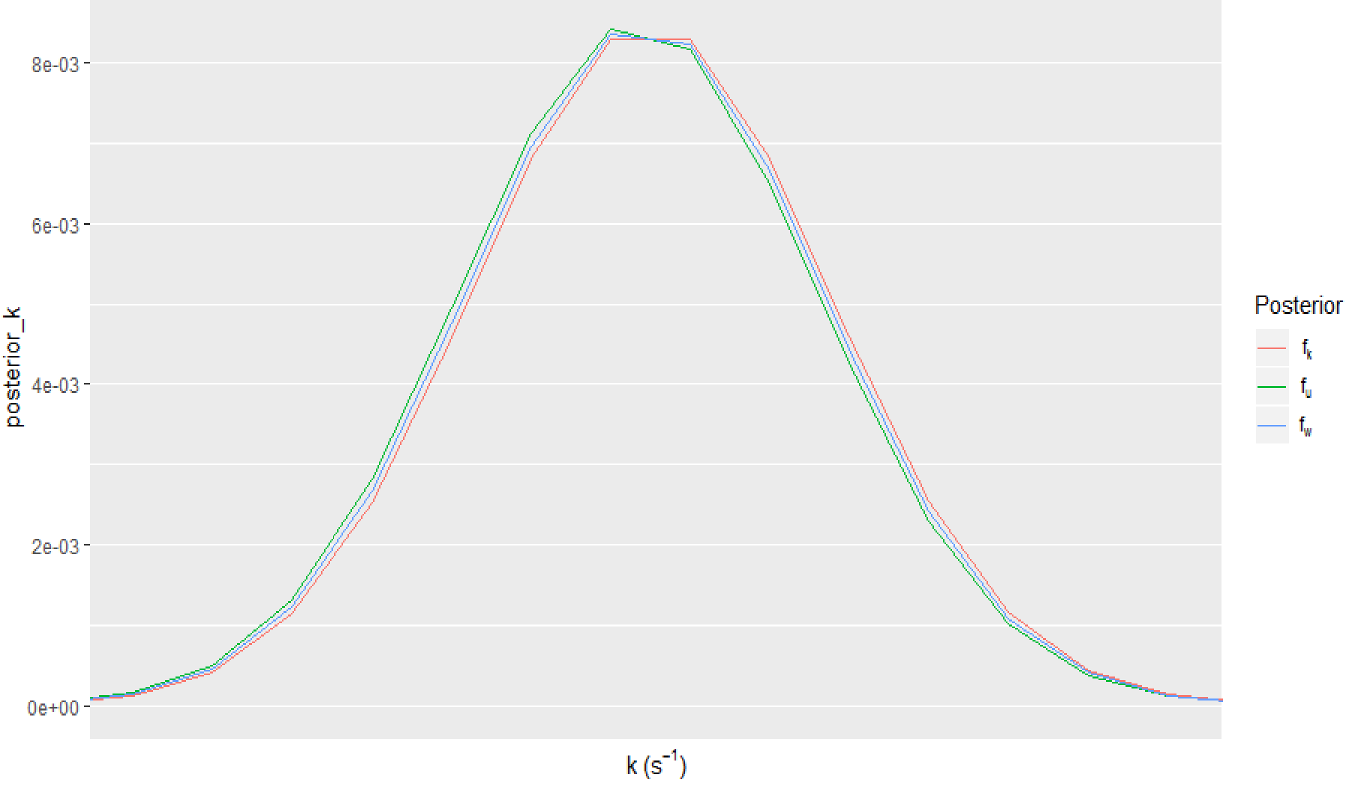
By plotting the posterior probability distributions corresponding to the three priors, we notice that their integrals is equal to 1.

```
[1] "Posterior integrals: "
[2] "1"
[3] "1"
[4] "1"
```

I. First Model

Question 5 :

By choosing an appropriate interval where the values of the likelihood are the highest, we obtain the following plot for the three posterior probability distributions.



I. First Model

Question 6 :

For each of the three priors, we compute the prior probability that *k* belongs to the interval [1000;1500].

```
lbound=2
ubound=4.2E+3
n_k=500
k = generate_grid_M1(lbound,ubound,n_k)

# #f_k
f_k_k <-  f_k(k,lbound,ubound)
integrate_density_M1(k, f_k_k, 1E+3, 1.5E+3)

#f_u
f_u_k <-  f_u(k,lbound,ubound)
integrate_density_M1(k, f_u_k, 1E+3, 1.5E+3)

#f_w
f_w_k <-  f_w(k,lbound,ubound)
integrate_density_M1(k, f_w_k, 1E+3, 1.5E+3)
```

We obtain the following result:

```
> # #f_k
> f_k_k <-  f_k(k,lbound,ubound)
> integrate_density_M1(k, f_k_k, 1E+3, 1.5E+3)
[1] 0.1123041
>
> #f_u
> f_u_k <-  f_u(k,lbound,ubound)
> integrate_density_M1(k, f_u_k, 1E+3, 1.5E+3)
[1] 0.0006311403
>
> #f_w
> f_w_k <-  f_w(k,lbound,ubound)
> integrate_density_M1(k, f_w_k, 1E+3, 1.5E+3)
[1] 0.05010216
```

No, the three prior probabilities are not in relatively good agreement with one another.

I. First Model

Question 7 :

The three corresponding posterior probabilities $p(k1 \in [1000\ ;1500] | experiment)$

```
R0= 1200
R_profile_exp = read.table("R_Exp.csv", header=FALSE)
lbound=2
ubound=4.26E+3
n_k=500
eps = 0.065
k = generate_grid_M1(lbound,ubound,n_k)
Compute_likelihood_all_M1(R0,lbound,ubound,n_k,eps)
L_all <- read.table("Likelihood_M1.csv", header=FALSE)
colnames(L_all) <- c("k_all","L_all")
L_all <- L_all$L_all
#f_k
f_k_k =  f_k(k,lbound,ubound)
post_k = Compute_posterior_M1(k, f_k_k, L_all)
I_k <- integrate_density_M1(k, post_k, 1E+3, 1.5E+3)
#f_u
f_u_k <-  f_u(k,lbound,ubound)
post_u = Compute_posterior_M1(k, f_u_k, L_all)
I_u <- integrate_density_M1(k, post_u, 1E+3, 1.5E+3)
#f_w
f_w_k <-  f_w(k,lbound,ubound)
post_w = Compute_posterior_M1(k, f_w_k, L_all)
I_w <- integrate_density_M1(k, post_w, 1E+3, 1.5E+3)

int = c(I_k,I_u,I_w)
print(c("The three corresponding posterior probabilities are: "))
print(int)
```

We obtain the following result:

```
[1] "The three corresponding posterior probabilities are: "
> print(int)
[1] 0.9973258 0.9978419 0.9975964
```

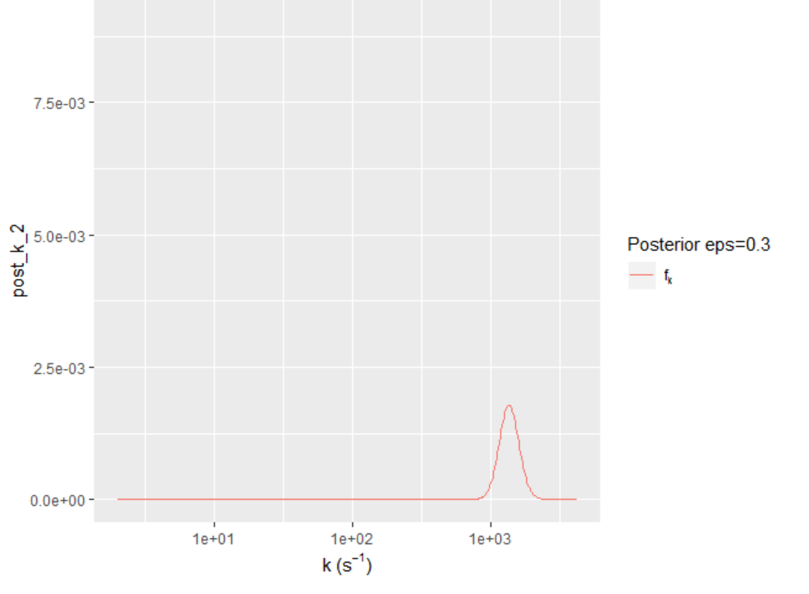
In this case, we obtain the same results for the three priors, so we conclude that the prior have a negligible effect on the posterior probability distribution.

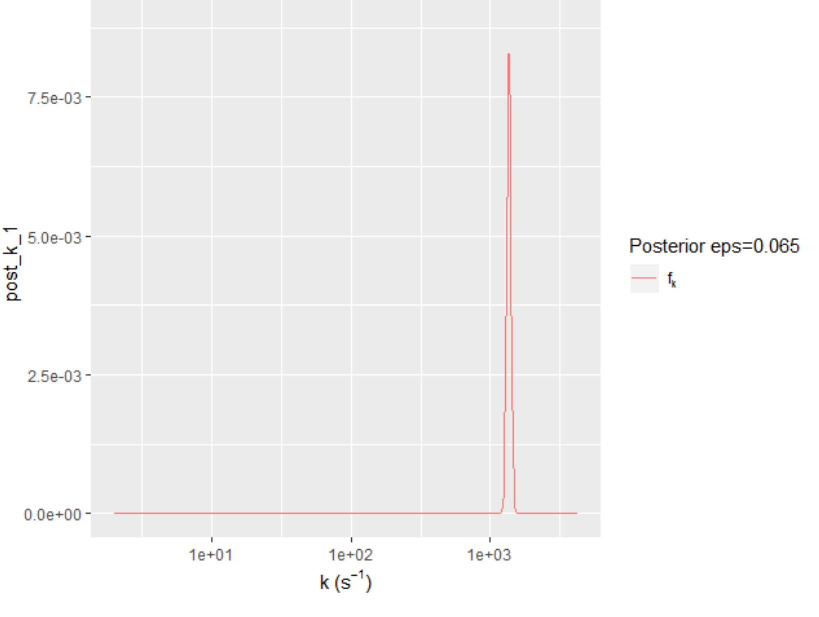
=> We say that the data wash out the prior.

I. First Model

Question 8 :

The posterior for ϵ=6.5 % and the posterior for ϵ=30% beside each other :





I. First Model

Question 9 :

The three corresponding posterior probabilities *p(k1∈[1000 ;1500]|experiment)*

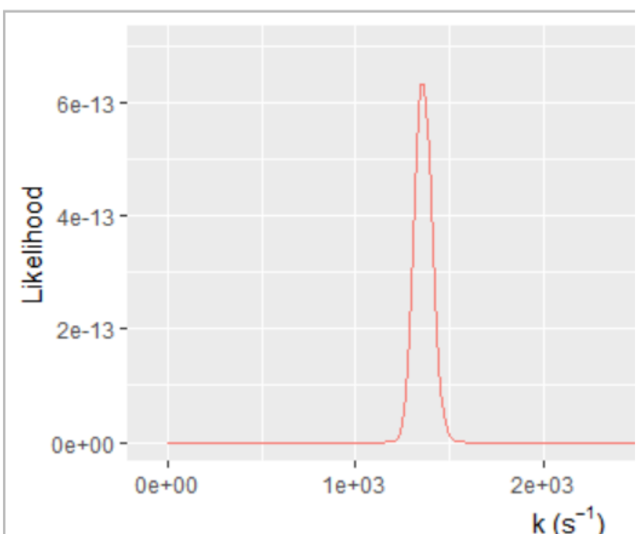
For epsilon = 0.065

```
[1] "If epsilon=0.065, the three corresponding posterior probabilities are: "
> print(int_1)
[1] 0.9973258 0.9978419 0.9975964
```

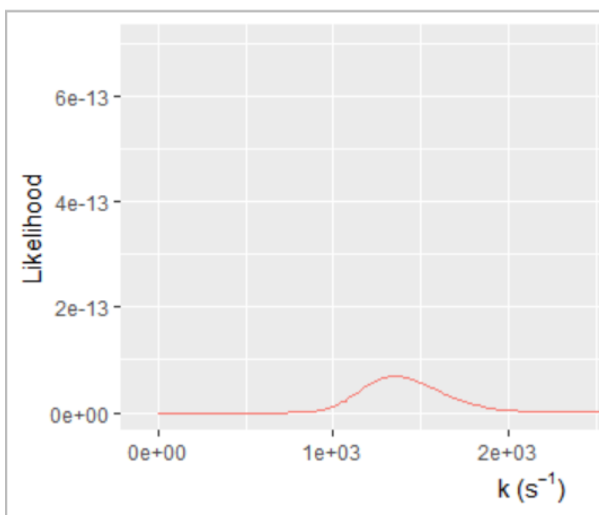
For epsilon = 0.3

```
> #eps2
> int_2 = c(I_k_2, I_u_2, I_w_2)
> print(c("If epsilon=0.3,The three corresponding posterior probabilities are: "))
[1] "If epsilon=0.3,The three corresponding posterior probabilities are: "
> print(int_2)
[1] 0.6313570 0.7130516 0.6768512
```

To explain the difference, we plot the likelihood for both epsilon = 0.065 and epsilon = 0.3



Likelihood for epsilon = 0.065



Likelihood for epsilon = 0.3

Knowing that The likelihood function intervenes in the value of posterior probabiliy. And according to the plot, we see that the maximum likelihood for esp= 0.065 is bigger than the esp = 0.3.

=> So, now we undesrtand the difference between the values of posterior probabilities.

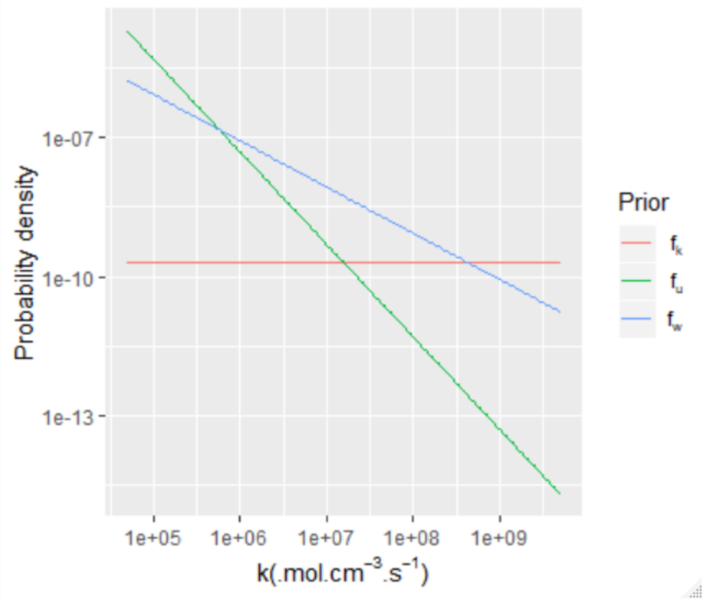
II. Second Model

Question 2 :

Using the following code:

```
lbound=5E+04
ubound=5E+09
n_k = 200
test_produce_priors_M2(lbound,ubound,n_k)
```

And by plotting the three priors, we obtain:



To verify that all is good, we calculate the integral and verify that it is equal to 1:

```
[1] "Value of the 3 integrals : " "1"                    "1.00167400352758"
[4] "1.0005579389276"
```

Good results !

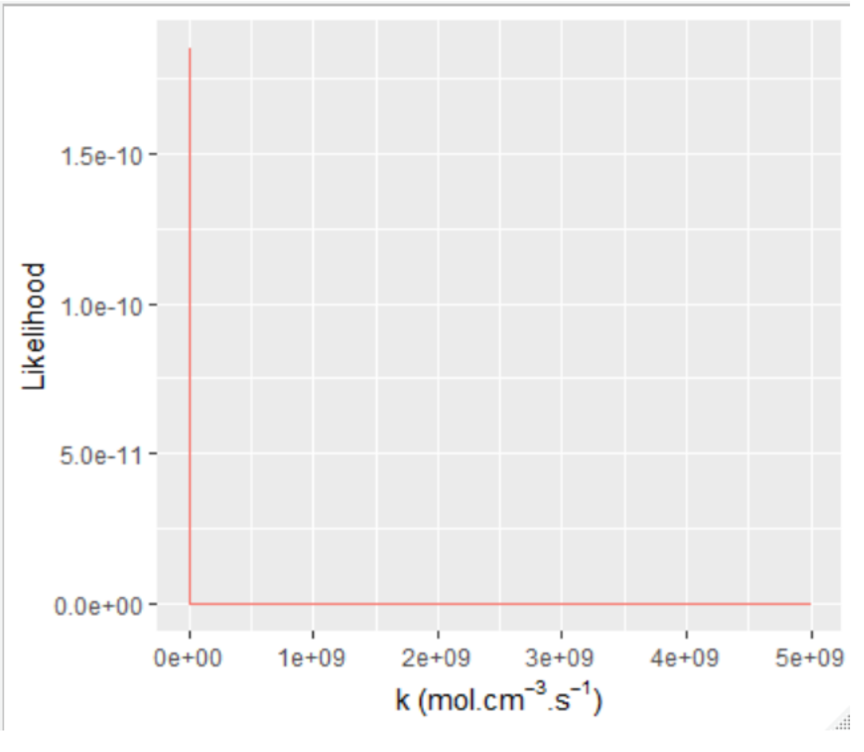
II. Second Model

Question 3 :

Using the following code:

```
R0= 1200
lbound=5E+04
ubound=5E+09
n_k=500
Examine_likelihood_M2(R0,lbound,ubound,n_k)
```

And by plotting the likelihood, we obtain:



As the maximum likelihood is located on k= 1003633.651

=> We choose to narrow the interval using the new one: [5 E+05; 1.5E+06]

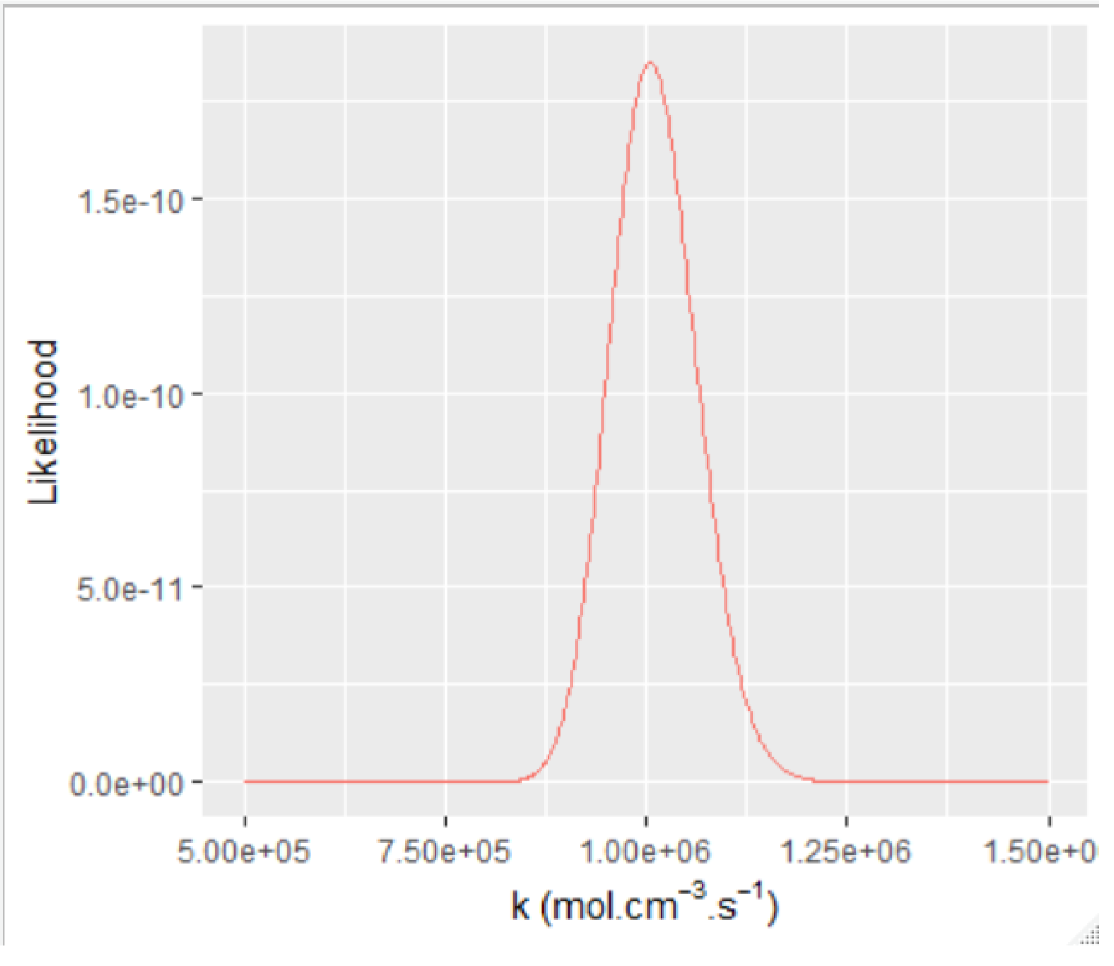
II. Second Model

Question 3 :

Using the following code, we focus on the specified interval:

```
R0= 1200
lbound=5E+05
ubound=1.5E+06
n_k=500
Examine_likelihood_M2(R0,lbound,ubound,n_k)
```

We obtain:



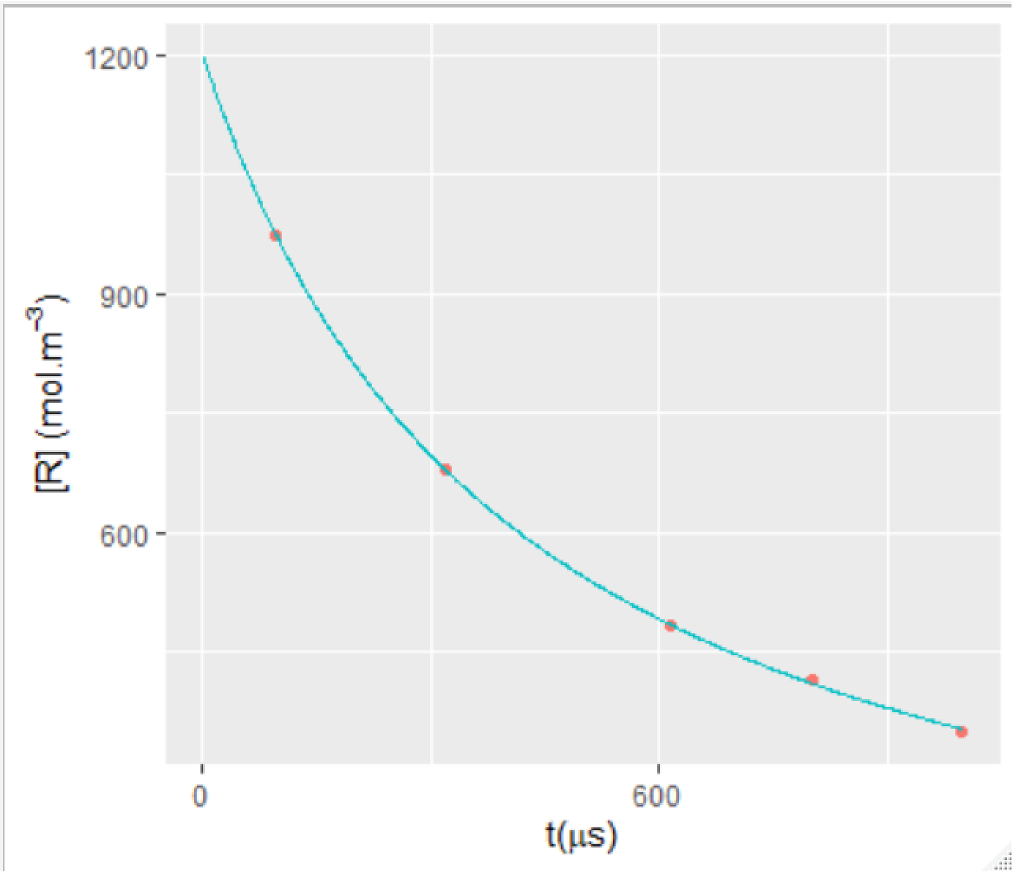
II. Second Model

Question 4 :

The optimal profil for the model M2 correspond to k=1.00363E+06.
Using the following code:

```
R0= 1200
k=1.003633E+06
t_mod=  seq(1,1000,1)
R_mod= Compute_R_profile_M2(t_mod,R0,k)
plot_profile_M2(t_mod,R_mod)
```

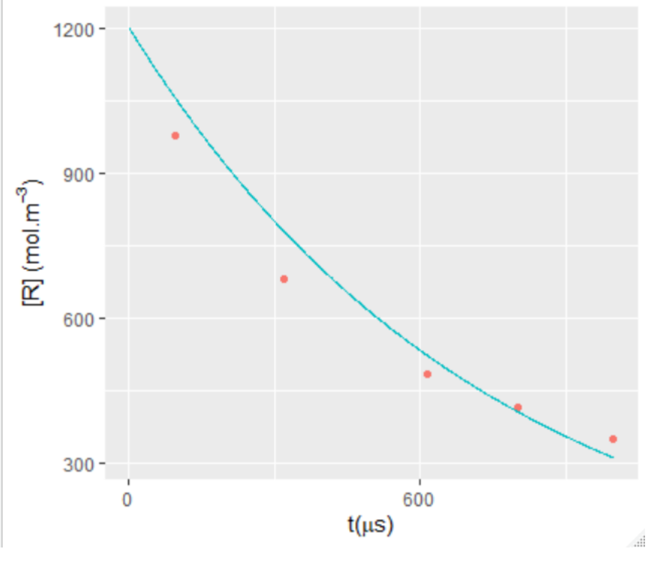
We obtain the following plot:



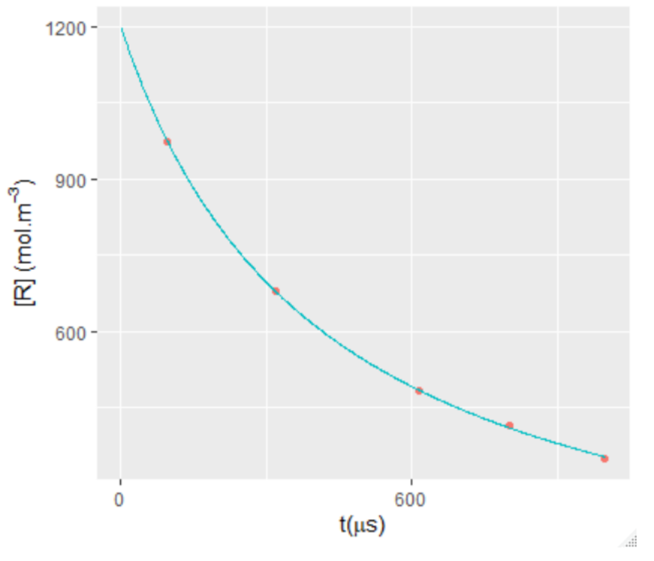
II. Second Model

Question 4 :

By comparing this profil with the optimal profil for M1, we notice that the model M2 is more appropriate since the curve goes through the experimental points.



Optimal profil for Model M1



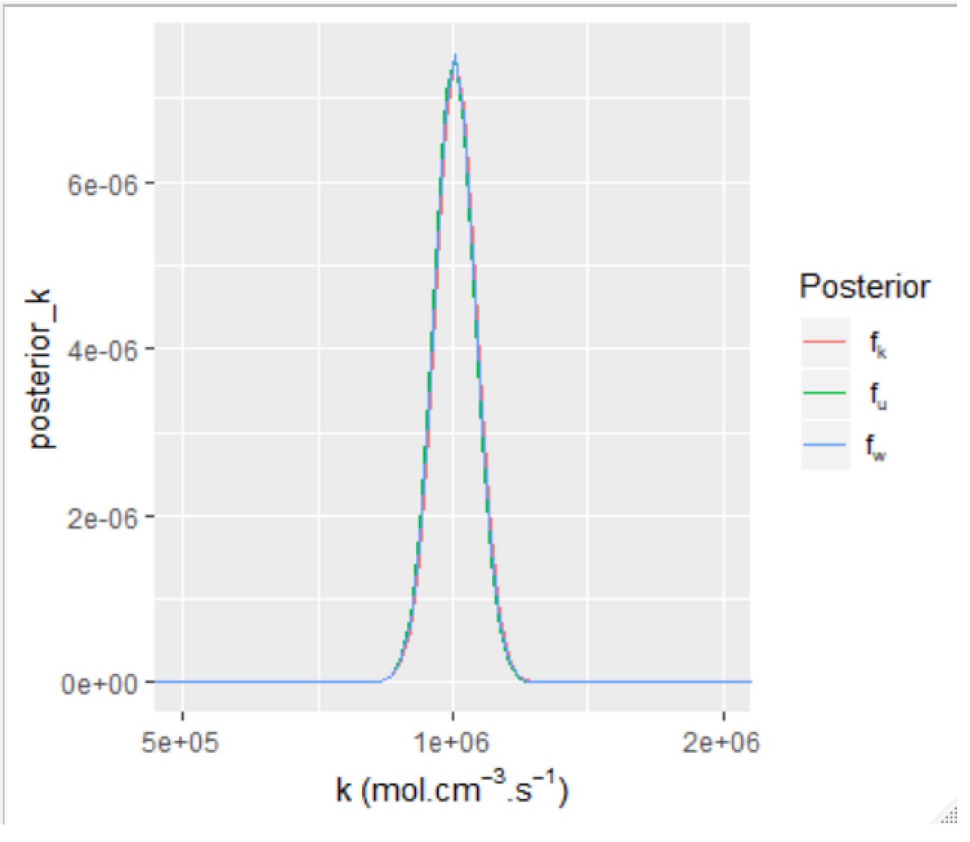
Optimal profil for Model M2

So model 2 better represents the experience.

II. Second Model

Question 5 :

We plot the three posterior probability distributions using an appropriate interval ([5 E+05; 1.5E+06]) where the values of the likelihood are the highest.



By plotting the three posterior probability distributions corresponding to the three priors, we notice that the 3 values of the returned vector are equals to ~1 :

```
[1] "Posterior integrals: " "1"                    "1"
[4] "1"
```

III. Comparison of model M1 and model M2

Question 1 : Calculate Bk

To calculate the value of the Bayes factor for the first uniform prior k, we use the following code:

```
R0=1200
lbound=2
ubound=4.26E+3
n_k=500
k = generate_grid_M1(lbound,ubound,n_k)
prior_M1=produce_all_priors_M1(lbound,ubound,n_k)
f_k_k = prior_M1[,2]
Compute_likelihood_all_M1 (R0,lbound,ubound,n_k,epsilon)
L <- read.table("Likelihood_M1.csv", header=FALSE)
colnames(L) <- c("k_all","L_all")
L_all <- L$L_all
B_1=Compute_B(k,f_k_k,L_all)


lbound=5E+4
ubound=5E+9
n_k=500
k = generate_grid_M2(lbound,ubound,n_k)
prior_M2=produce_all_priors_M2(lbound,ubound,n_k)
f_k_k = prior_M2[,2]
Compute_likelihood_all_M2 (R0,lbound,ubound,n_k,epsilon)
L <- read.table("Likelihood_M2.csv", header=FALSE)
colnames(L) <- c("k_all","L_all")
L_all <- L$L_all
B_2=Compute_B(k,f_k_k,L_all)

paste("Bk= ",B_2 / B_1)
```

By executing this code, we obtain the Bk value:

```
> paste("Bk= ",B_2 / B_1)
[1] "Bk=  0.276463385455728"
```

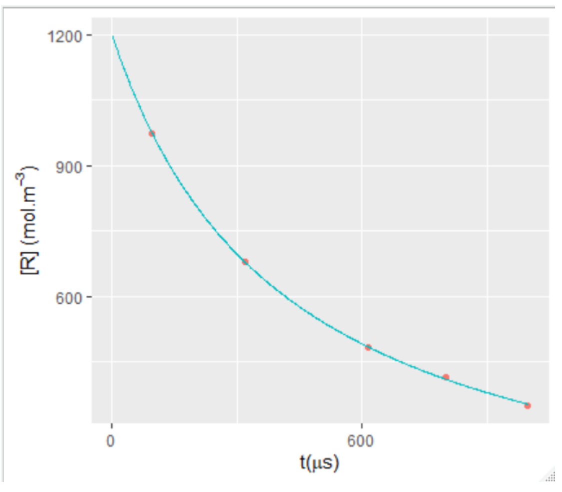
This value mean that the data disfavour M2 and support M1.

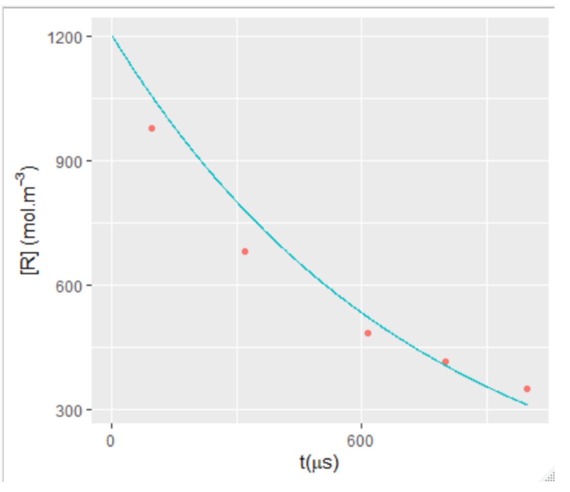
III. Comparison of model M1 and model M2

Question 1 : Calculate Bk

Commentary:

Considering the plot of the optimal profil of M1 and M2, **we notice a paradoxe.** Because normaly B should favour M2.





=> To explain this, we calculte Bu and Bw.

III. Comparison of model M1 and model M2

Question 2 : Calculate Bu & Bw

By doing the same steps as in question 1, we found that:

For Bu:

```
> paste("Bu= ",B_2 / B_1)
[1] "Bu=  14751.3932751854"
```

And for Bw:

```
> paste("Bw= ",B_2 / B_1)
[1] "Bw=  291.17913089373"
```

From this results, the values of Bu and Bw mean that the data strongly favour M2. Which is real considering the plot of the optimal profil of both models