Stochastic modeling of the LacI system with multiple feedback using Langevin approach

The following report is arranged as follows:-

1. **Introduction**
2. **Description of the 4 different strains developed**
3. **Effect of IPTG on system**
4. **Effect of lactose**
5. **A brief overview of stochasticity in biological processes**
6. **Langevin approach- a brief explanation**
7. **The model equations for each of the 4 strains**
8. **Modelling of growth**
9. **Solution Strategy**
10. **Key results and discussions**
11. **Appendix**

Introduction

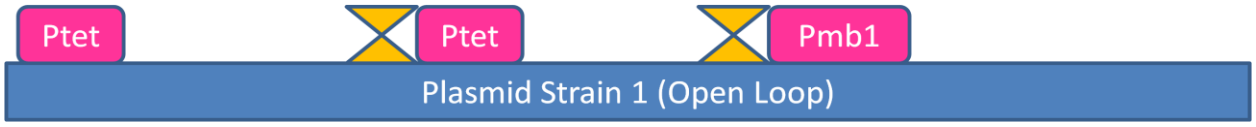
We wish to demonstrate the utility of a multiple feedback control system as compared to a single feedback control system. This utility is characterized in terms of faster response with reduced noise.

We would be using the lac operon system in an E. coli cell for this purpose. The lac operon system is a complex mechanism used for the digestion of lactose. The lac operon system when expressed results in the formation of the protein lacI. We have introduced a modified plasmid in the E. coli cell. The lacI produced from the operon within this plasmid has a cfp(cyan fluorescent protein) attached to it, whose fluorescence is an indicator of the amount of lacI present in the system. Further, the plasmid number also keeps growing with a separate mechanism. The plasmid is attached with a yfp(yellow fluorescent protein) so that the fluorescence of the protein acts as an indicator of its number. The promoter of the lacI system(ptet) can further be modified(to plac) such that lacI can itself bind to the promoter and repress its expression. This forms one level of feedback where the process is being regulated by the output itself. Further the replication of plasmid number can be modified in such a way that the lacI produced inhibits the process. Thus we have a unique structure where the process output can regulate not only the process itself, but also the number of processes.

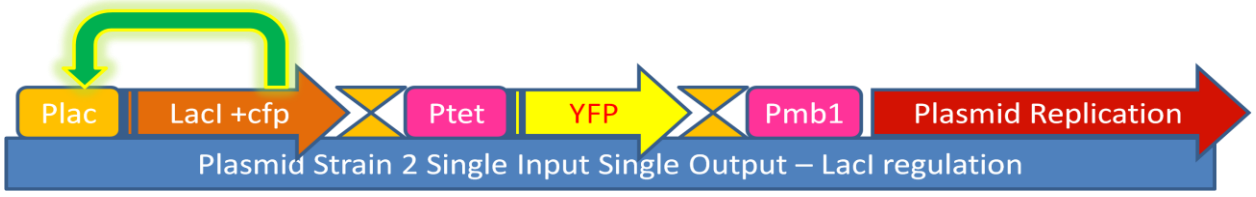
Thus we have 2 control levels. By combination, we have 4 different control loops or structures possible, expressed in 4 different strains.

Description of the 4 strains:-

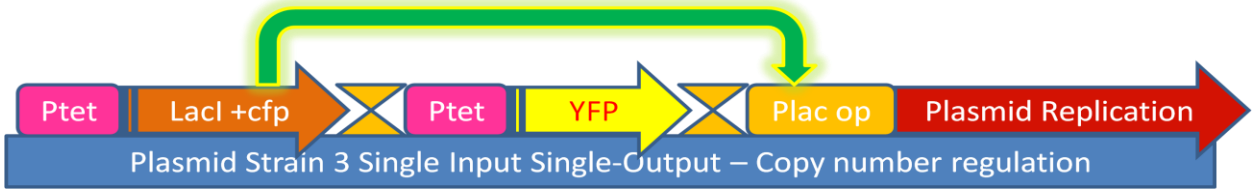
Strain-1 (Open loop):- **The promoter of both lacI and plasmid number replication is unaffected by lacI. This forms the open loop or the simplest of structures which is unregulated.**



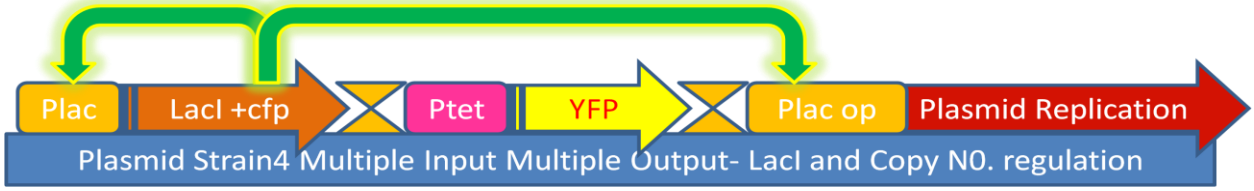
Strain-2 (SISO_lacI):- **The promoter for the lac system is changed to plac. Thus lacI can bind to its promoter site and prevent its expression. This forms one level of regulation. The copy number is unregulated.**



Strain-3(SISO_CN):- **The promoter for replication of plasmid is modified such that lacI can bind to its site and prevent replication. Thus the copy number is regulated. LacI expression is unregulated.**



Strain-4(MIMO):- **Both the promoters are modified so that lacI produced can regulate the expression of lacI as well as the replication of plasmids. This forms the highest level of control.**



Effect of IPTG on system

Further, an interesting study can be conducted by addition of IPTG in the system. IPTG(isopropyl-beta-D-thiogalactopyranoside), is an inducer for the lac system. It binds with lacI. Thus in any of the regulated strains IPTG can potentially bind to any of the lacI molecules which are attached to the promoter sites, releasing them and thus causing expression of the system. Thus a regulated system with a high amount of IPTG should resemble with the open loop system in its behavior. We wish to characterize this effect too.

Effect of Lactose

The lacI operon in the presence of lactose leads to the production of the enzyme B-galactosidase which is required for the utilization of lactose. The cell produces the enzyme only when lactose is present in the system. Hence, when lactose is present in the system, the enzyme will be synthesized causing the biomass to increase. The growth so obtained would also be affected by the level of feedback existing within the cell. LacI can bind to the lactose, and it thus represents a burden on the cell. The growth that can be possible achieved with lactose thus competes with the lacI present which acts as a burden.

It is this growth v/s burden effect that we wish to characterize on addition of lactose.

Stochasticity in Biological Processes

A stochastic process is the counterpart of a deterministic process, where, instead of assuming that the system with the same initial conditions will always give the same determined response, we allow for indeterminacy in the future response of a system. Thus there is always some random or noise element present, which implies that there are quite a few possible paths or trajectories for a system, though a few might be more probable than others.

Stochasticity in biological systems has been well established . Stochasticity for any system is an inverse function of square root of number of particles. Biochemical species participating in processes such as transcription, translation often do so in very low numbers, and thus intrinsically, biochemical processes always encounter a lot of noise. Characterization of this noise is therefore extremely essential.

A detailed deterministic model analysis of the system has already been attempted. We now wish to characterize the inherent noise or stochasticity of the system, and attempt to demonstrate the reduction in the same in the regulated strains.

Langevin Approach:

The approach that we apply to introduce stochasticity in our model is simple yet powerful. As detailed above, all biochemical processes have inherent stochasticity. We introduce this stochasticity in the system by allowing for a range of values for the kinetic parameters. Thus instead of having a fixed set of values for all parameters, we perturb them within a certain range, and then characterize the deviations caused from the mean output. This simple method can also give us detailed insights about the responses and their differences in all the 4 strains.

Further, since the effect of perturbations or random elements is the key part we want to identify through our modeling, we simplify the detailed model to ignore all the facilitated diffusion and other effects. We assume that the internal concentrations of proteins like lacI, IPTG within the cell are same as external concentrations.

Model Equations for each of the strains

Strain -1 Open Loop:

We just consider the simple differential changes for the lacI protein and copy number.

The protein expression depends on the copy number. Further, the protein concentration would be affected by dilution and degradation.

Similarly, the copy number replication would depend on the existing copy number, and would be affected by both dilution and degradation. The copy number replication can be expected to follow a Michelis Menton type of kinetics. Further, since the open loop system is unaffected by IPTG, we see no dependence for the same herein.

Thus, the differential equations for the system would be:-

$$\frac{dlacI}{dt} = K_1 * Cn - (\mu + \beta)Cn$$

$$\frac{dCn}{dt} = K_2 * \frac{Cn}{K_c + Cn^2} - (\mu + \beta_2) * Cn$$

Where Cn is copynumber.

The parameters K_1 and K_2 are the parameters perturbed.

Strain-2 SISO_LacI :

The open loop mechanics for the system is slightly modified for the strain with lacI regulation. The lacI produced tends to repress the expression for the lacI operon. This effect can be represented by a controller term C1 in the generation equation for lacI as given below. The formation of the complex between IPTG and lacI is a fast step that reaches equilibrium. This equilibrium step is represented by a controller term C3, since the IPTG also acts as a switch for the system. We are interested in the total amount of lacI in the system.

$$\frac{dlacI}{dt} = K_1 * Cn * C1 - (\mu + \beta)Cn - K_4 * C3$$

$$\frac{dCn}{dt} = K_2 * \frac{Cn}{K_c + Cn^2} - (\mu + \beta_2) * Cn$$

$$\frac{d(lacI - IPTG)}{dt} = K_4 * C3$$

Where the terms C1 and C3 representing control action are:-

$$C1 = \frac{{k_1}^2}{{k_1}^2 + LacI^2}$$

$$C3 = \frac{IPTG * LacI^2}{{k_4}^2 + IPTG^2}$$

Strain -3 SISO_CN:

When only the copynumber is being regulated by lacI, the IPTG can still complex with lacI but it cannot regulate lacI expression. Thus from the above case, the controller C3 is present but C1 is absent. Further since copynumber is regulated by lacI, this regulation can be expressed in terms of a controller term C2, which is expressed as follows:-

$$\frac{dlacI}{dt} = K_1 * Cn - (\mu + \beta)Cn - K_4 * C3$$

$$\frac{dCn}{dt} = K_2 * \frac{Cn * C2}{K_c + Cn^2} - (\mu + \beta_2) * Cn$$

$$\frac{d(lacI - IPTG)}{dt} = K_4 * C3$$

Where the terms C2 and C3 representing control action are:-

$$C2 = \frac{{k_2}^2}{{k_2}^2 + LacI^2}$$

$$C3 = \frac{IPTG * LacI^2}{{k_4}^2 + IPTG^2}$$

STRain-4 (MIMO):-

This is the fourth strain with the highest level of control. All the controller terms, C1, C2, C3 discussed in the two strains above are present in this case. Hence the equations for this strain are:-

$$\frac{dlacI}{dt} = K_1 * Cn * C1 - (\mu + \beta)Cn - K_4 * C3$$

$$\frac{dCn}{dt} = K_2 * \frac{Cn * C2}{K_c + Cn^2} - (\mu + \beta_2) * Cn$$

$$\frac{d(lacI - IPTG)}{dt} = K_4 * C3$$

Where the terms C2 and C3 representing control action are:-

$$C1 = \frac{{k_1}^2}{{k_1}^2 + LacI^2}$$

$$C2 = \frac{{k_2}^2}{{k_2}^2 + LacI^2}$$

$$C3 = \frac{IPTG * LacI^2}{{k_4}^2 + IPTG^2}$$

Noise is introduced in this system by allowing all the parametric kinetic coefficients(all the k's) to be perturbed in a range of 30% off their mean values. The noise of many such runs is calculated by simulation.

The values of the parameters used are summarized:-

K1=50 min^{-1} K2=0.0045 $\mu M^2/l^2$-min k_1=1000 µM/l

k_2=1500 µM/l k_4=100 µM/l Kc=25 $\mu M^2/l^2$

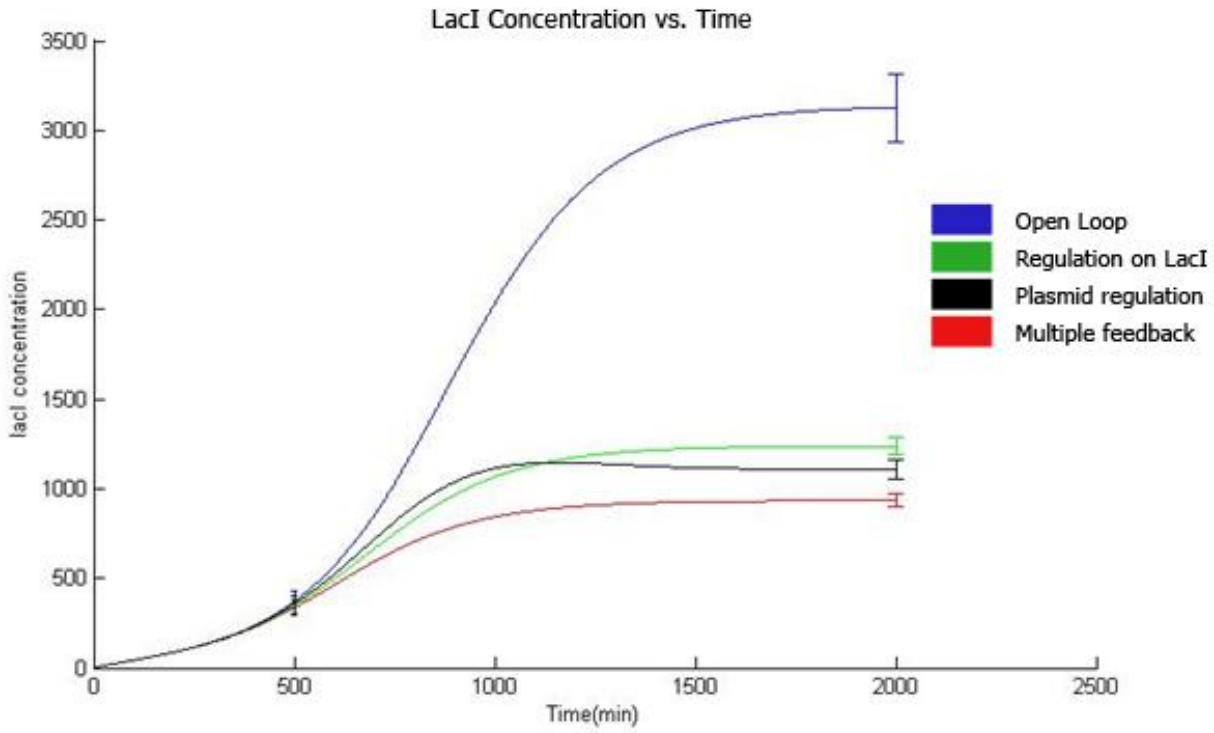
µ=0.4/60 min^{-1} β1=0.1/60 min^{-1} β2=0.01*β1 min^{-1}

Solution Strategy:-

Each of the models is solved in Matlab using in built differential equation solvers. The values of all the kinetic parameters and the control parameters(viz K1,K2,k1,k2,k4,Kc) are randomly perturbed from their mean values to a maximum limit of 30 % of their mean values. 100 simulations are performed for the same and their mean values are obtained to give a mean or an expected profile. The profiles so obtained for each of the 4 strains are plotted. Further, at each time instant, we will have a distribution of values for lacI and copynumber, from the 100 runs so performed. We fit a normal to these distributions and obtain the standard error at early time instants(t=500min) and at steady state value(t=2000min) . It is the effect on standard error that we are most interested in for each of the 4 strains. Further, similar simulations are performed by varying the amount of IPTG in the system. We are interested in the steady state values of lacI and copynumber obtained by the system. The plot of IPTG concentration v/s LacI and copynumber concentrations are also plotted, and compared for each of the 4 strains, the hypothesis being that each of the strains resembles the open loop behavior for high IPTG values. Further, growth is characterized in the open loop strain and strain with multiple feedback. The results for the same are also summarized.

Results:-

The profile for lacI v/s time for each of the strains are plotted below.



Thus, qualitatively each of the 4 strains exhibit similar behavior. The open loop shows maximum expression as expected , while the strain with multiple feedback exhibits least expression, as it is the one with maximum control. The important aspect to be noted is the error bars. Although they are indistinguishable on for the four strains at time t=500min, one can clearly see the difference in the four strains as they reach steady state. Their values are tabulated below.

The values of mean for lacI and the standard errors for each of the 4 strains for 100 runs at time t=500 and at time t=2000 mins is summarized below:-

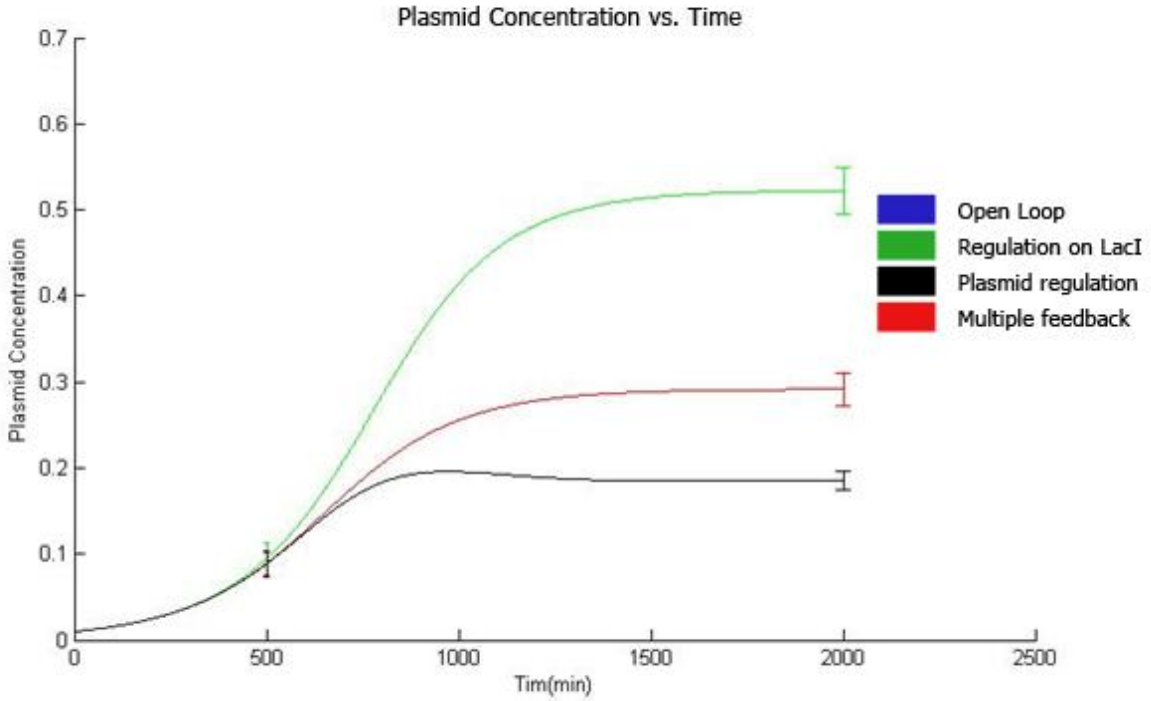
LacI

Strain	Mean(t=500)	Std Error(t=500)	Mean(t=2000)	Std Error(t=2000)
Open loop	458	32	2913	95
SISO_lacI	384	22	1182	24
SISO_CN	413	26	1059	27
MIMO	360	19.5	873	18

The mean value at steady state for lacI concentration is significantly reduced for the strain with multiple feedback. This can again be attributed to better control as compared to the open loop strain which shows unregulated and unrestrained expression.

Thus the error is almost one half for the strain with multiple feedback initially, which further becomes $1/5^{th}$ as the systems reach steady state. This validates our hypothesis that the strain with multiple feedback is more tightly regulated. Further, the strain with lacI regulation also exhibits better control than the strain with plasmid regulation which is as expected.

*See Appendix for all the required distributions and graphs



Thus, qualitatively each of the 4 strains exhibit similar behavior. The blue and the green lines coincide, the strain with lacI regulation and open loop show no difference for plasmid replication. The open loop shows maximum expression as expected , while the strain with plasmid regulation exhibits least expression. This decrease as compared to the strain with multiple feedback can be attributed to the fact that in the strain with multiple feedback, the lacI produced is utilized for the regulation of both lacI expression and copy number replication, and thus in the strain with plasmid regulation, there is more lacI present in the system for plasmid regulation only. The important aspect to be noted is the error bars. Although they are indistinguishable on for the four strains at time t=500min, one can clearly see the difference in the four strains as they reach steady state. Their values are tabulated below.

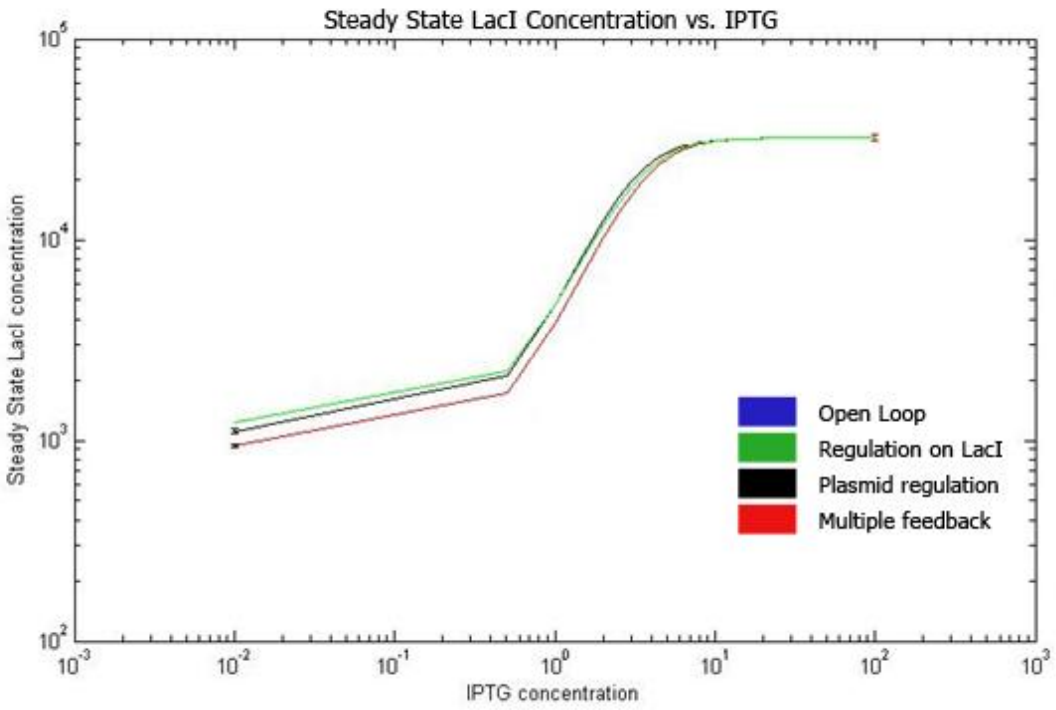
The means and errors for plasmid concentration are tabulated, at time t=500 min and at time t=2000 min

Plasmid Concentration

Strain	Mean(t=500)	Std Error(t=500)	Mean(t=2000)	Std Error(t=2000)
Open loop	0.125	0.0094	0.5	0.014
SISO_lacI	0.125	0.0094	0.5	0.014
SISO_CN	0.105	0.007	0.2	0.005
MIMO	0.11	0.0075	0.3	0.01

Here as well, the multiple feedback strain shows a great reduction in noise and expression as compared to the open loop. The strain with lacI regulation shows no difference from open loop, which is expected since plasmid replication is independent of lacI concentrations for both cases. The error is reduced by $1/3^{rd}$ for the case with plasmid regulation as compared to open loop, while the steady state values are also almost halved.

Effect of IPTG on system:-



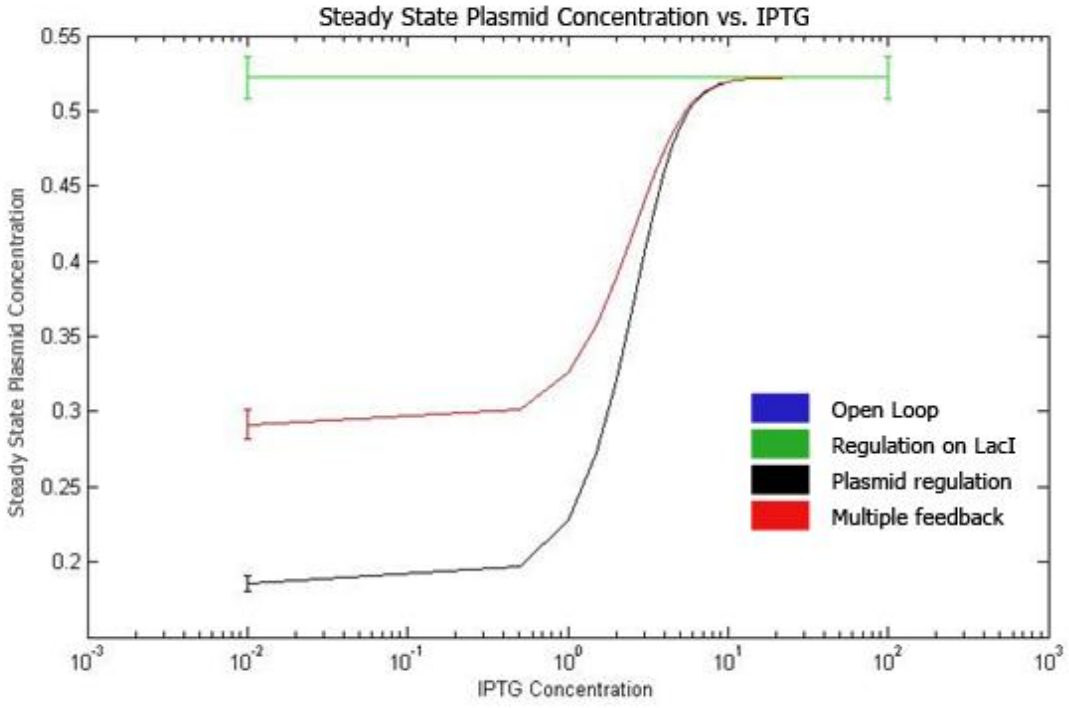
The steady state lacI concentration v/s IPTG profile is plotted above. As expected, all 4 strains show similar qualitative behavior, with differences pronounced at low IPTG concentrations. At higher IPTG concentrations, all strains exhibit behavior similar to that of open loop strain and hence the all 4 strains converge to the same steady state values. The open loop is unaffected by the presence of IPTG and hence it is not shown here.

The simulations are carried out 100 times for a low IPTG value(0.01) and high IPTG value(100) using the randomly perturbed parameters. The distributions so obtained and their fit are attached in the appendix. The results are summarized.

Steady State LacI concentration

Strain	Mean(IPTG=0.01)	Std Error(IPTG=0.01)	Mean(IPTG=100)	Std Error(IPTG=100)
Open loop	2913	95	2913	95
SISO_lacI	1182	24	29840	1355
SISO_CN	1059	27	29855	1357
MIMO	873	18	29822	1354

At lower IPTG values, differences are obtained in the 4 strains with the strain with multiple feedback exhibiting less standard error as explained previously. At higher IPTG values, all strains exhibit the same steady state means and errors. Note that the steady state concentration obtained is of the total lacI present in the system(Free+ complexed with IPTG). Hence the observed increase in the order of magnitude of steady state value of lacI at higher IPTG.



The steady state plasmid concentration v/s IPTG profile is plotted above. The strain with regulation on lacI is unaffected by IPTG as expected, while the other two exhibit similar qualitative behavior , with differences pronounced at low IPTG concentrations. At higher IPTG concentrations, all strains exhibit behavior similar to that of open loop strain and hence the all 4 strains converge to the same steady state values. The open loop is unaffected by the presence of IPTG and hence it is not shown here.

The simulations are carried out 100 times for a low IPTG value(0.01) and high IPTG value(100) using the randomly perturbed parameters. The distributions so obtained and their fit are attached in the appendix. The results are summarized.

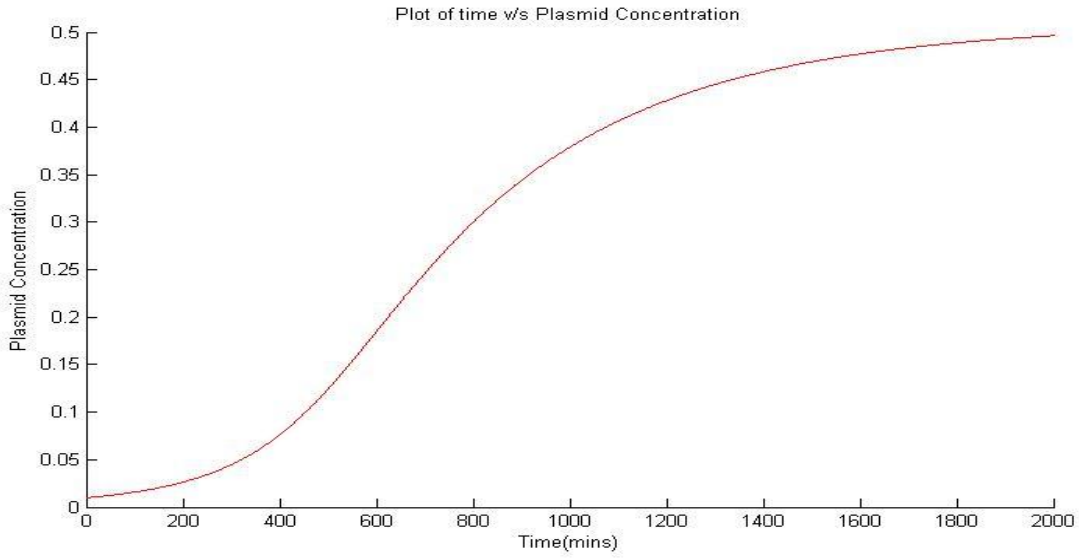
Steady State Plasmid Concentration

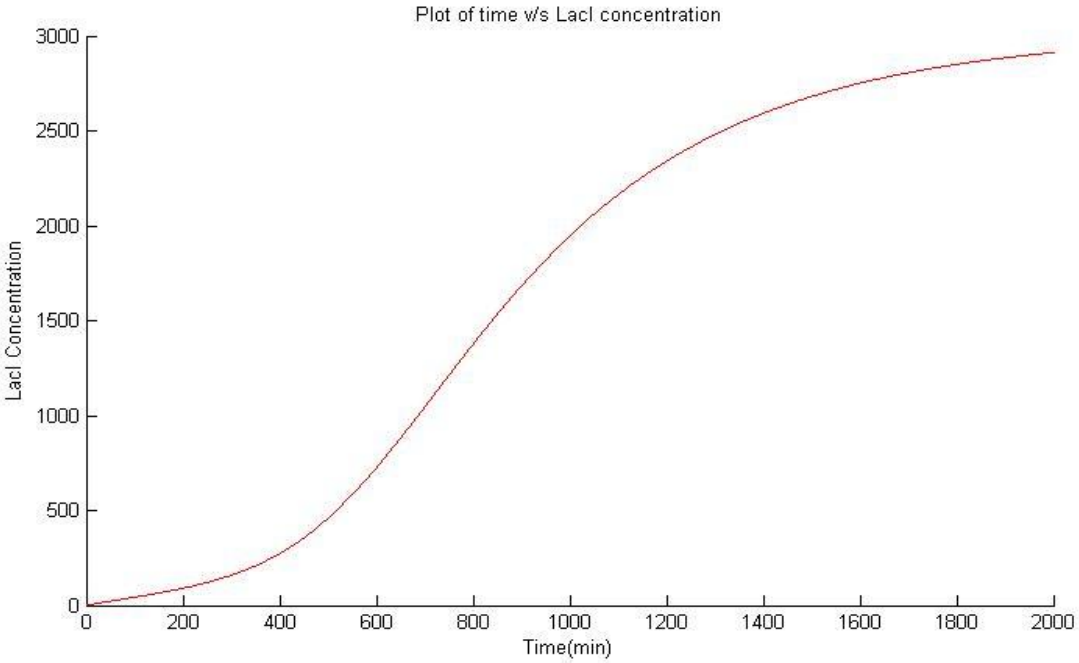
Strain	Mean(IPTG=0.01)	Std Error(IPTG=0.01)	Mean(IPTG=100)	Std Error(IPTG=100)
Open loop	0.5	0.014	0.5	0.014
lacI regulation	0.5	0.014	0.5	0.014
Plasmid Regulation	0.2	0.005	0.5	0.0138
Multiple Feedback	0.3	0.01	0.5	0.0138

Thus, even in this case, we observe the similarity in the behavior of the 4 strains with increasing IPTG concentrations.

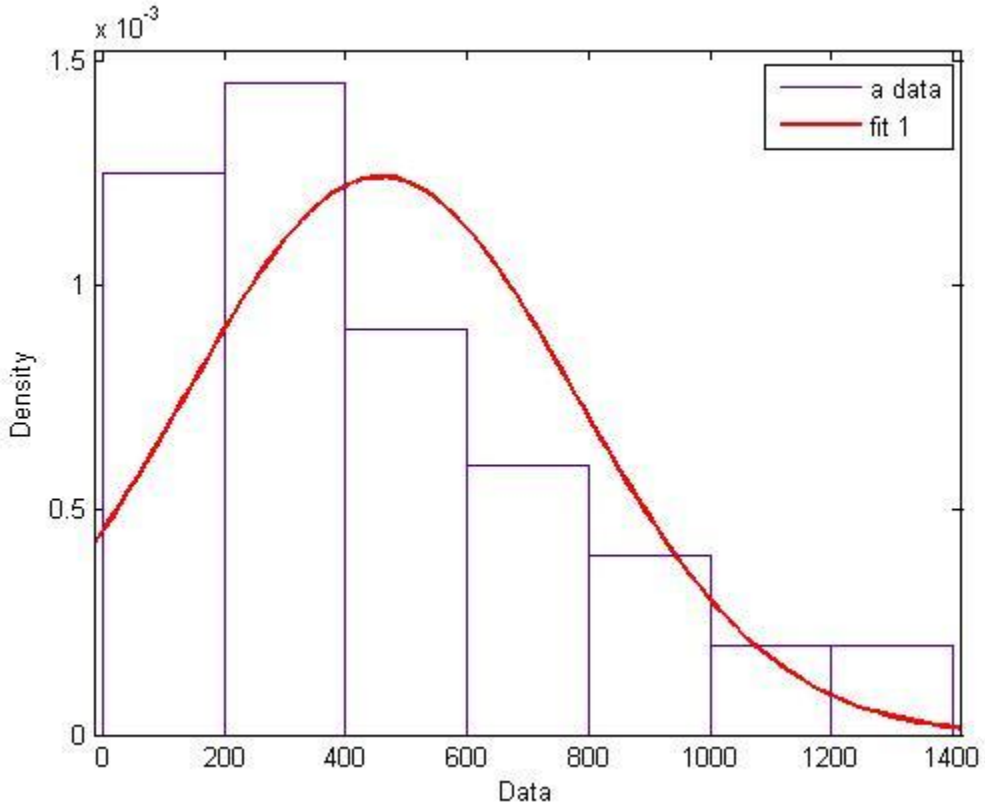
Appendix

For the Strain-1 the time domain profiles are characterized below:-





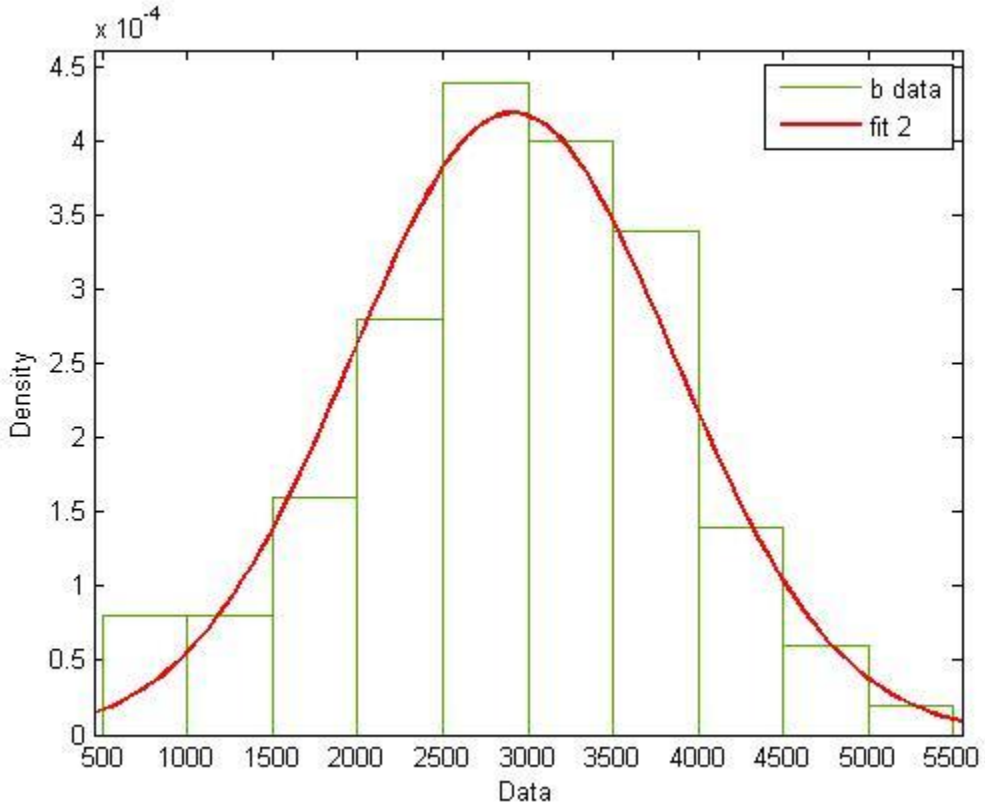
The distribution of lacI concentration at t=500 min obtained by 100 runs



mu 457.865

Std. Err. 32.1751

The distribution of lacI concentration at t=2000 min obtained by 100 runs

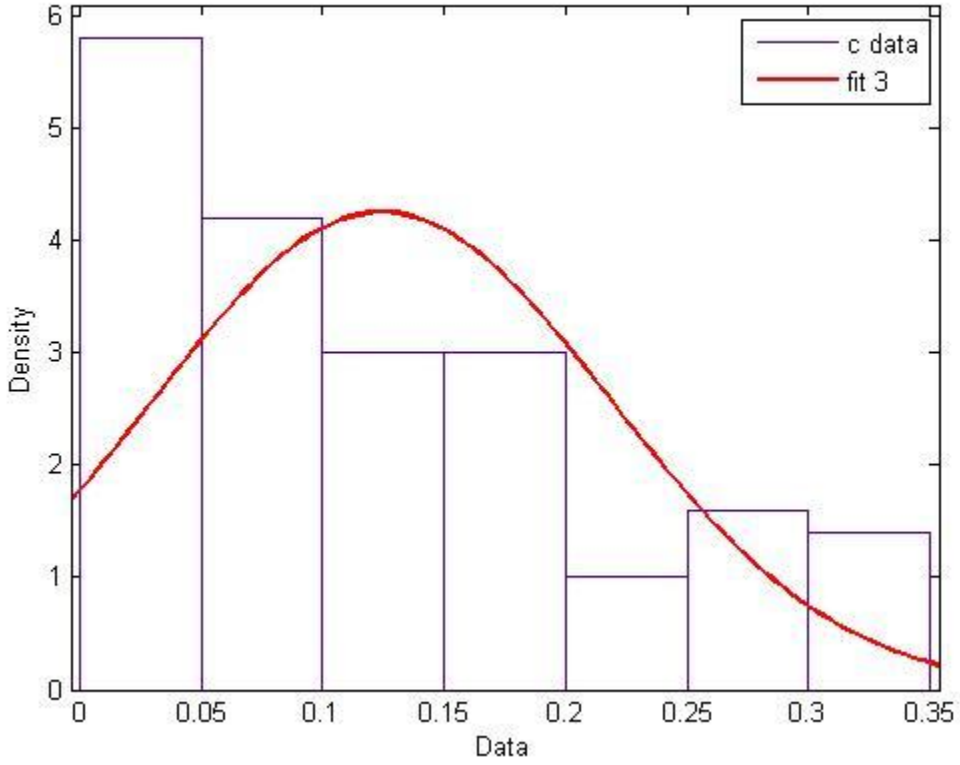


.

mu 2912.57

Std error 95.0793

The distribution of plasmid concentration at t=500 min obtained by 100 runs

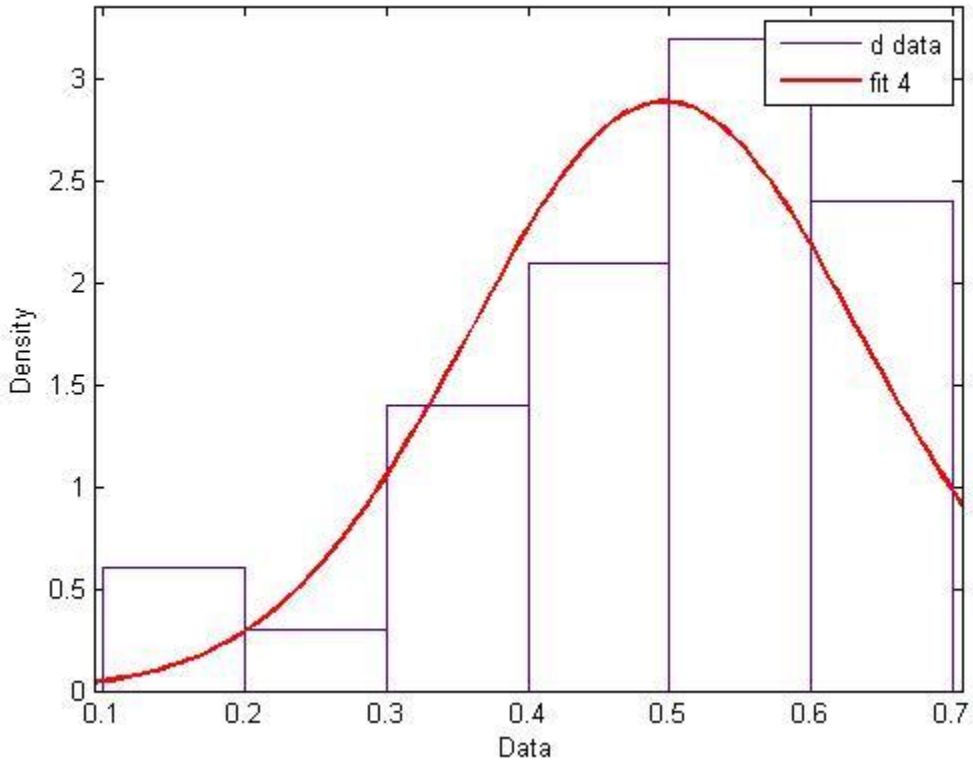


.

mu 0.124424

Std Error 0.0093915

The distribution of plasmid concentration obtained at t=2000 min by 100 runs

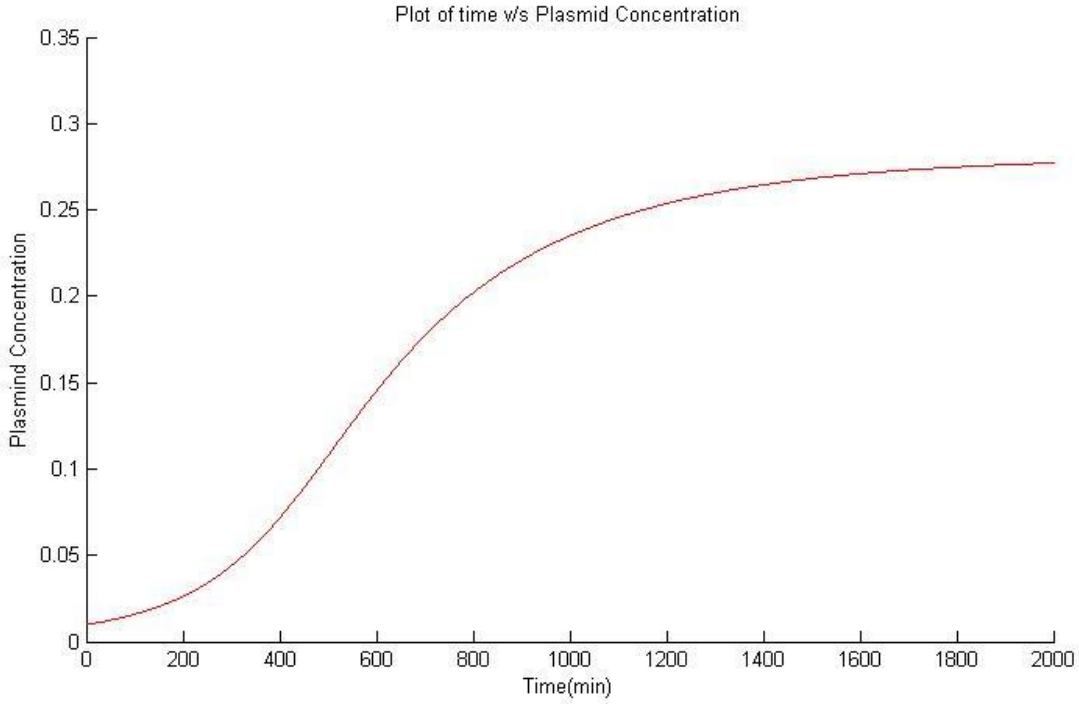


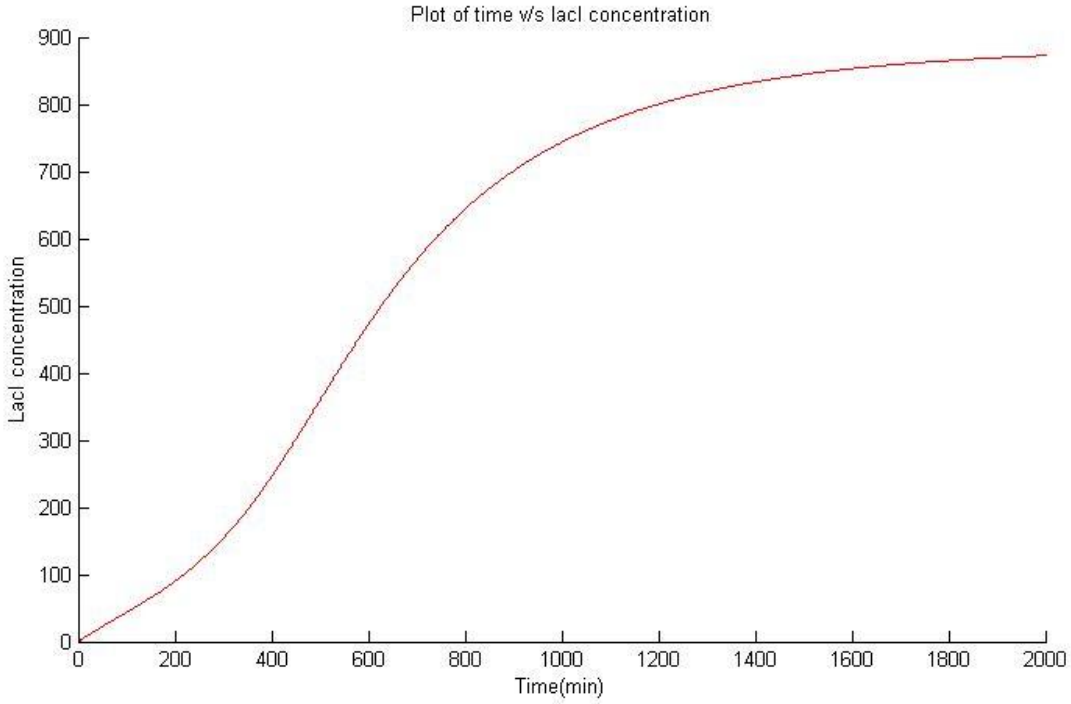
mu 0.496259

Std. Err.0.0138032

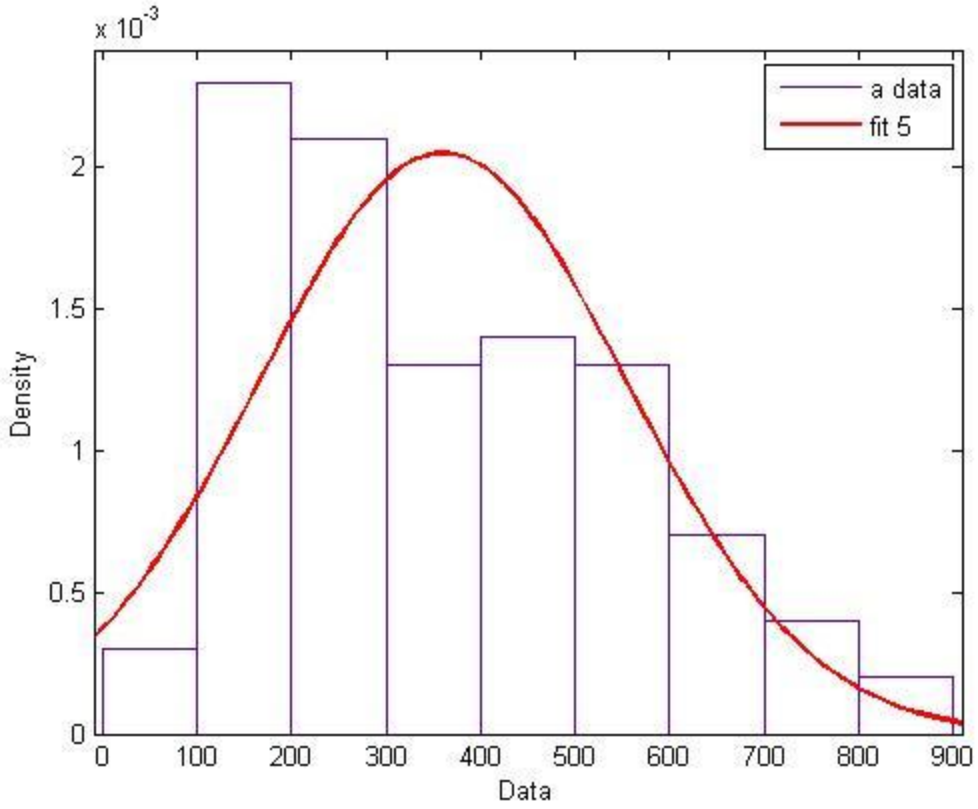
Strain-4 (MIMO):-

The response of the system v/s time is plotted below.





The distribution of lacI concentration obtained at t=500 min by 100 runs

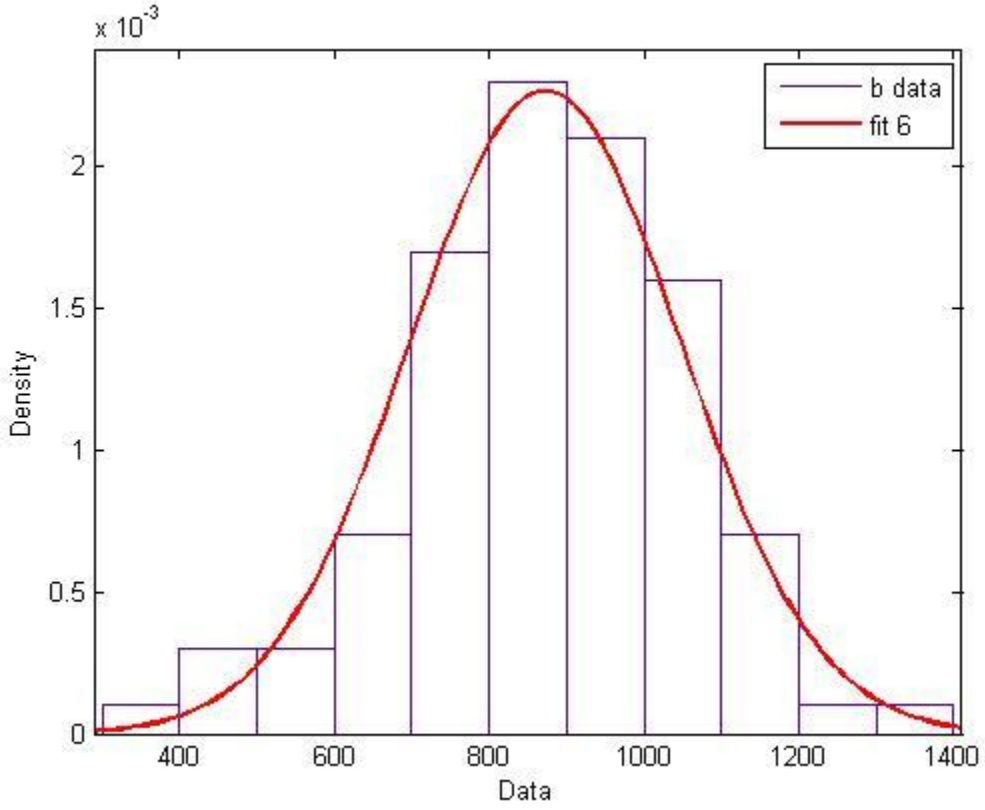


.

mu 360.034

Std Error 19.4673

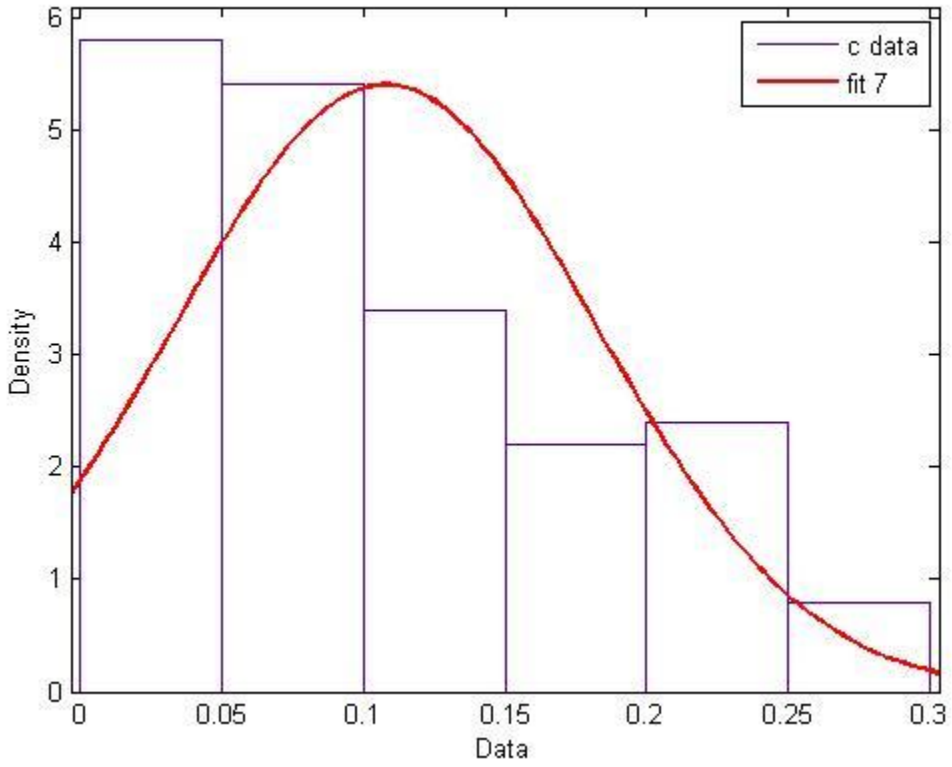
The distribution of lacI concentration at t=2000 min obtained by 100 runs



mu 872.65

Std Error 17.5968

The distribution of plasmid concentration obtained at t=500 min by 100 runs

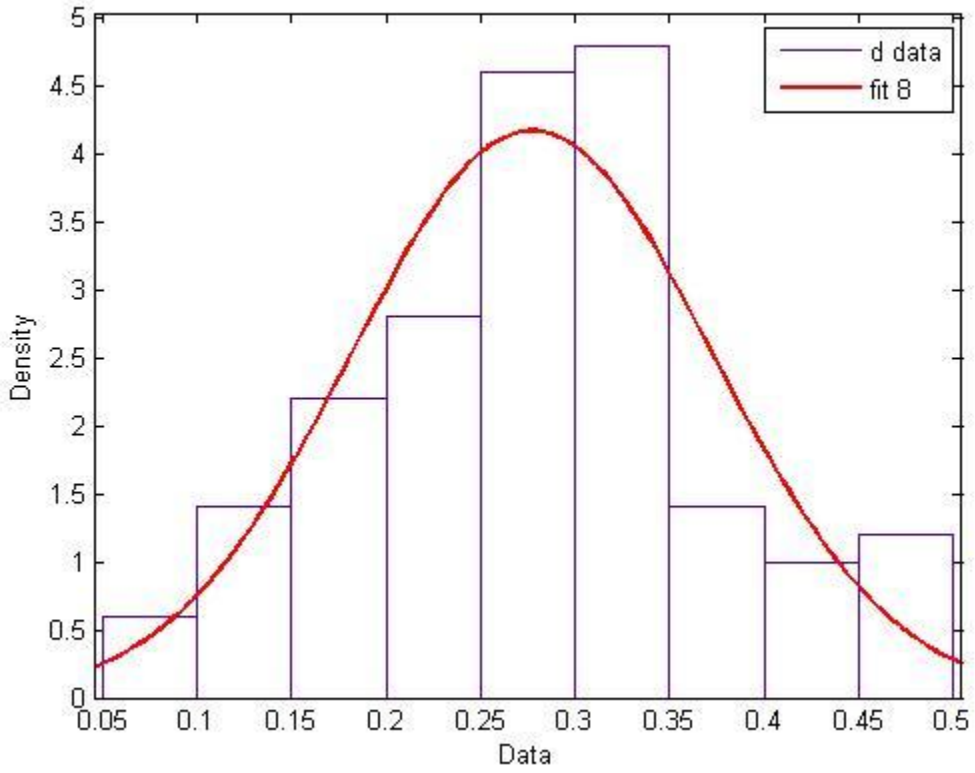


.

mu 0.107868

Std. Err 0.00739123

The distribution of plasmid concentration obtained at t=2000 min by 100 runs

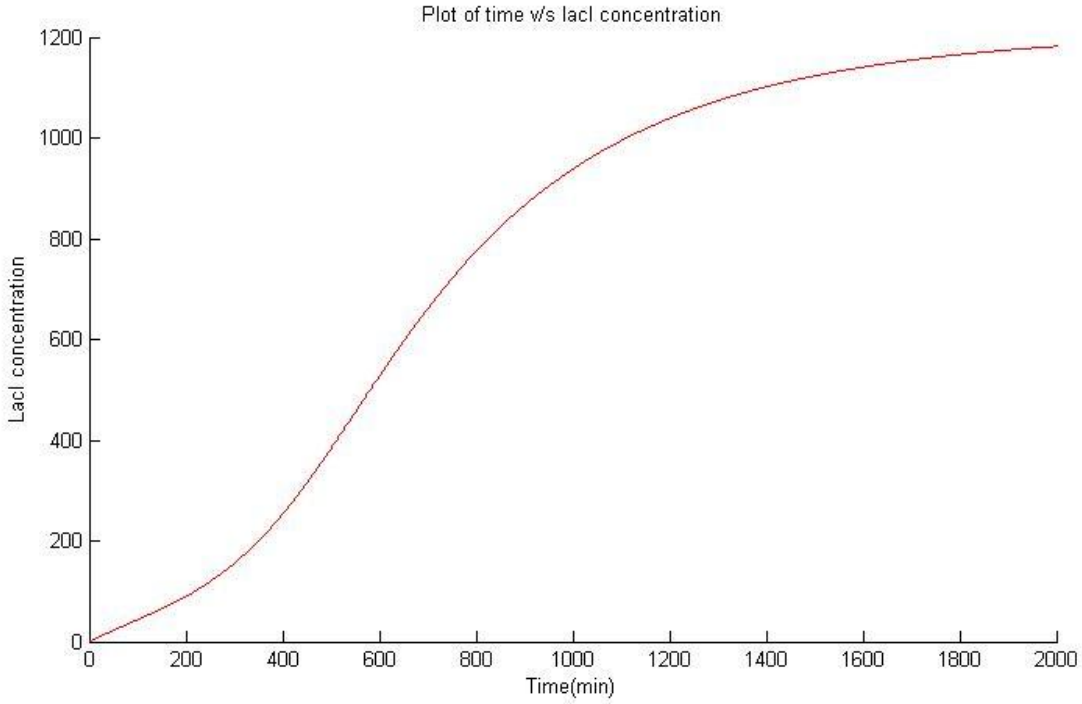


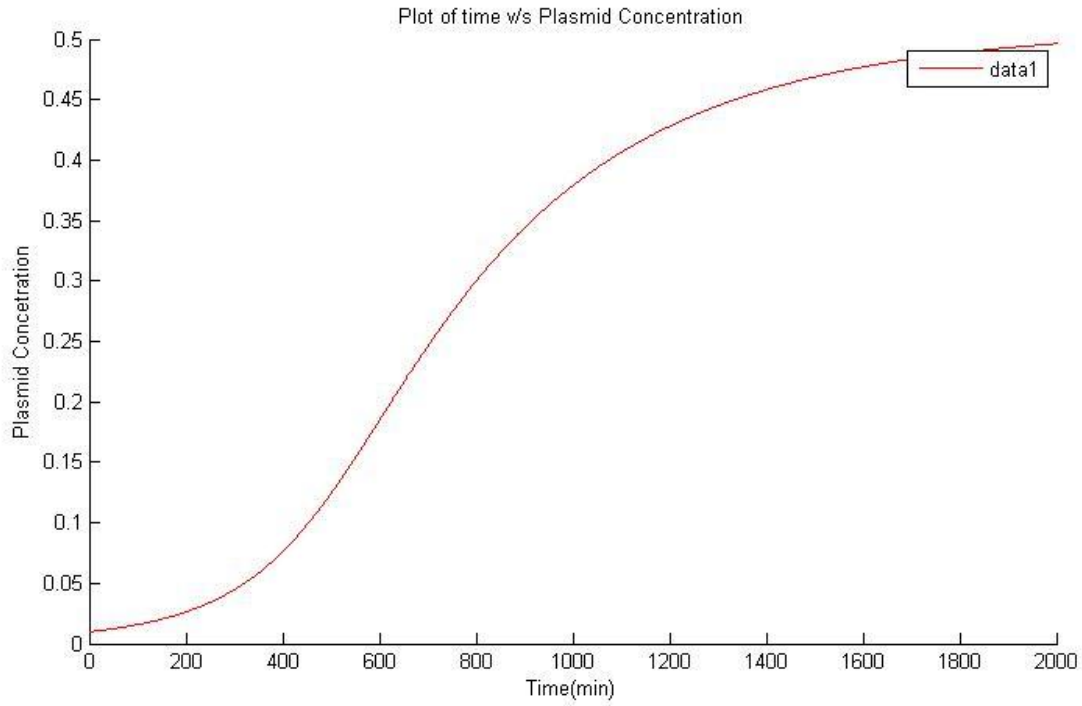
mu 0.277137

Std. Err. 0.00955624

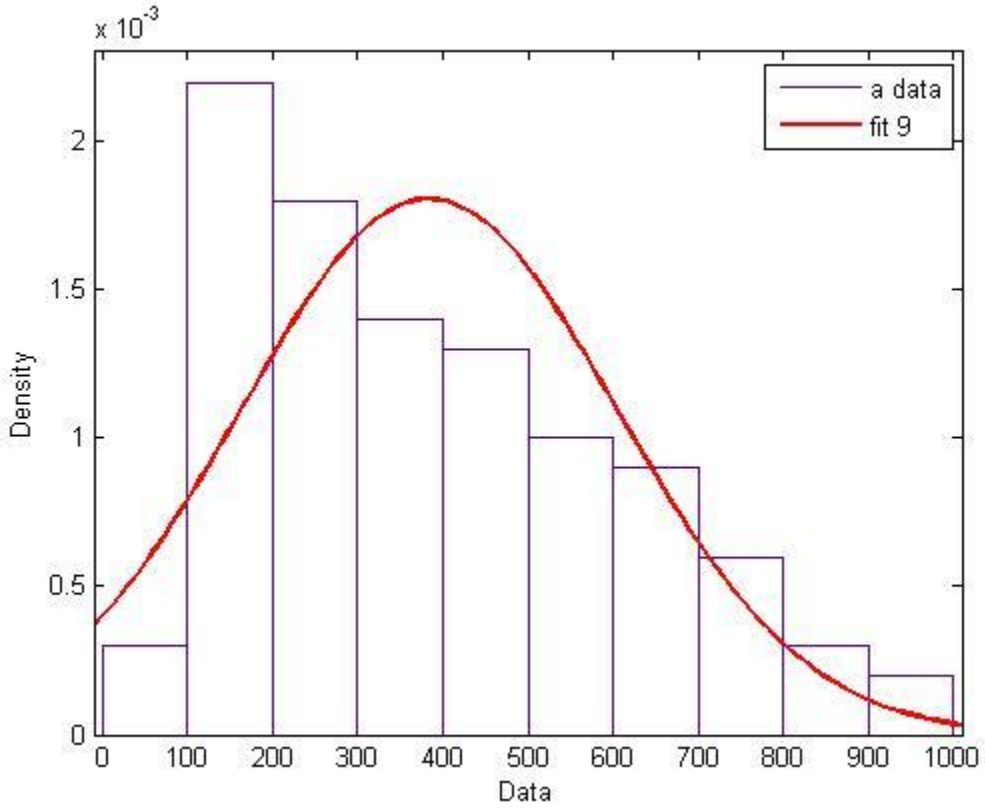
Strain-2(SISO_lacI):-

The responses v/s time of the strain with lacI regulation are plotted below





The distribution of lacI concentration at t=500 min obtained by 100 runs

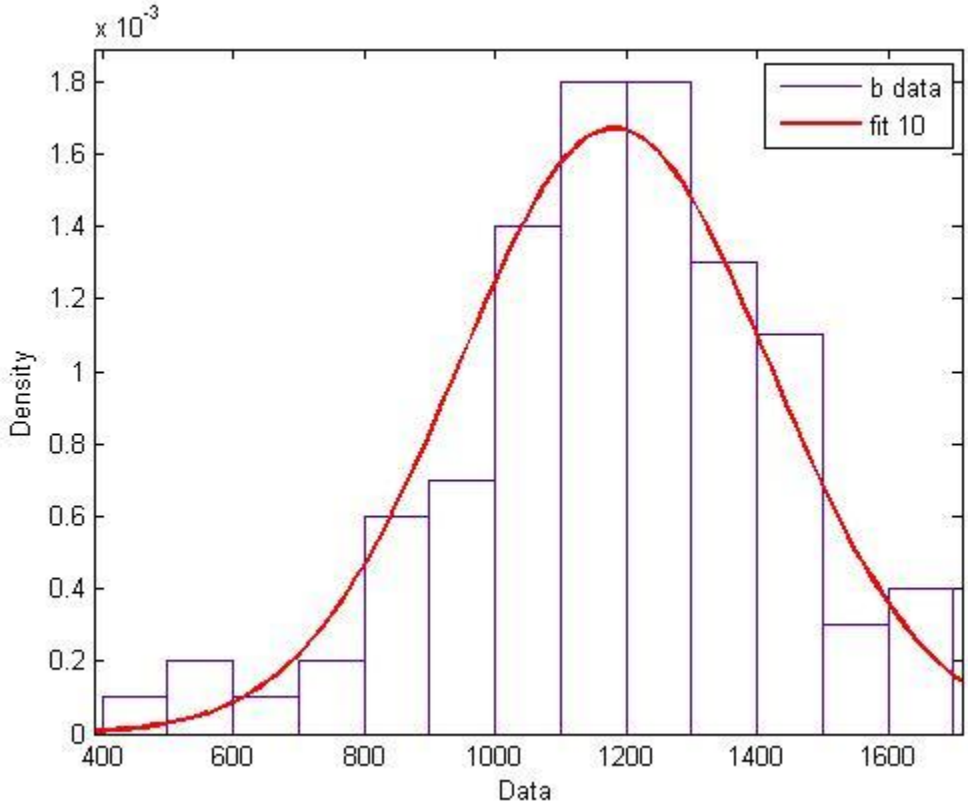


.

mu 383.837

Std. Err 22.0777

The distribution of lacI concentration at t=2000 min obtained by 100 runs

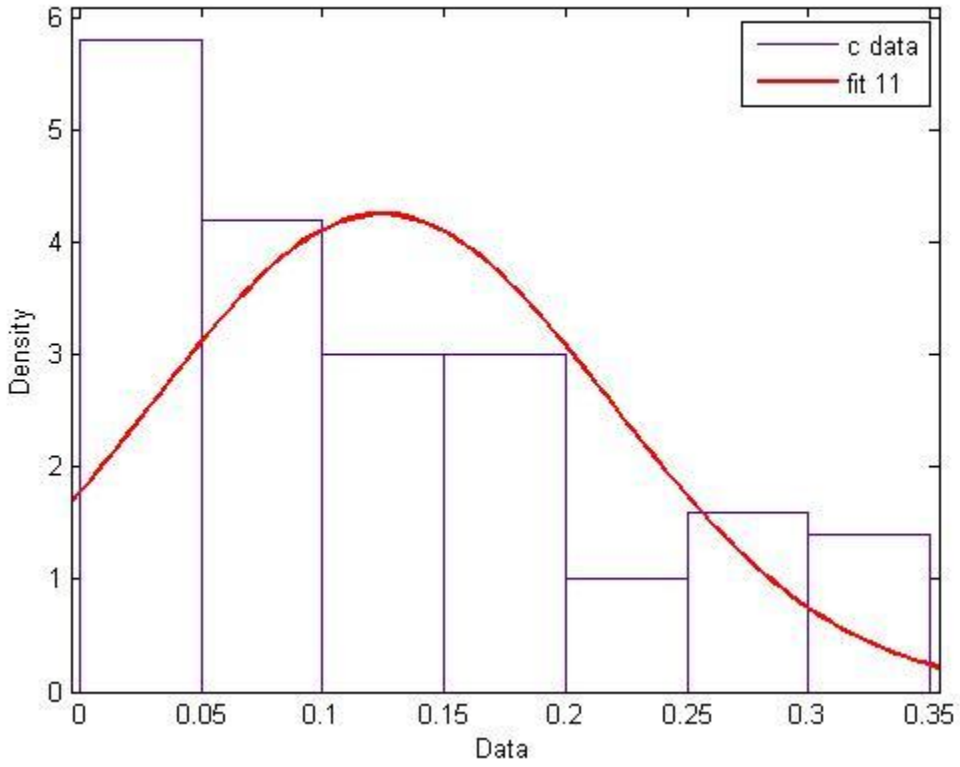


.

mu 1181.86

Std. Err 23.8863

The distribution of plasmid concentration at t=500 min obtained by 100 runs

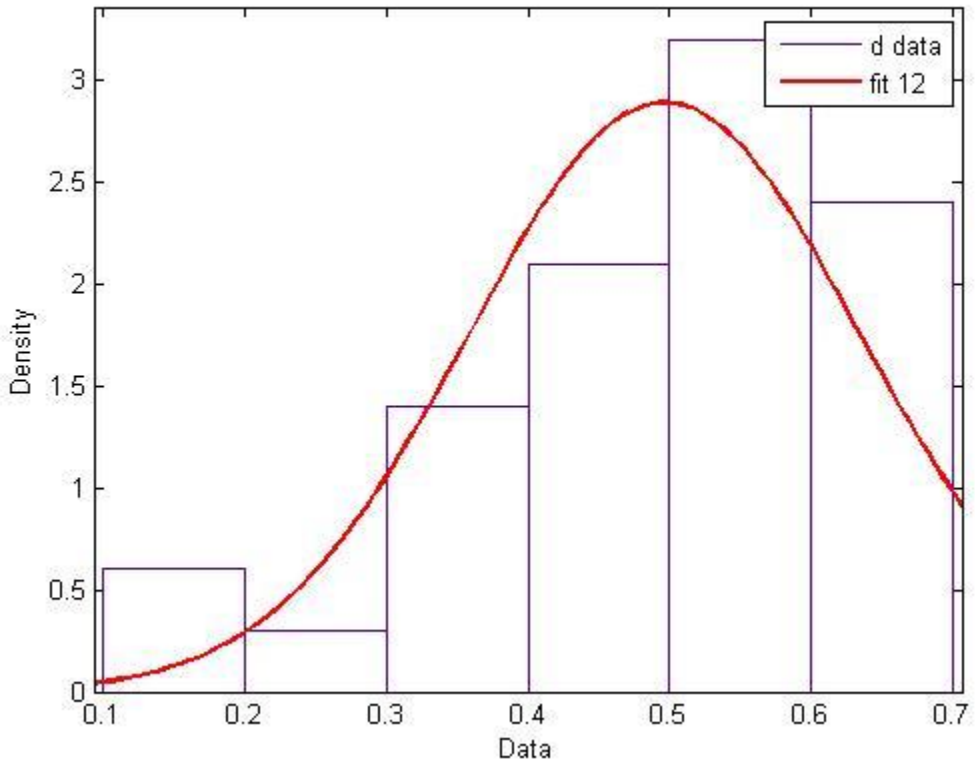


.

mu 0.124424

Std. Err 0.0093915

The distribution of plasmid concentration at t=2000 min obtained by 100 runs



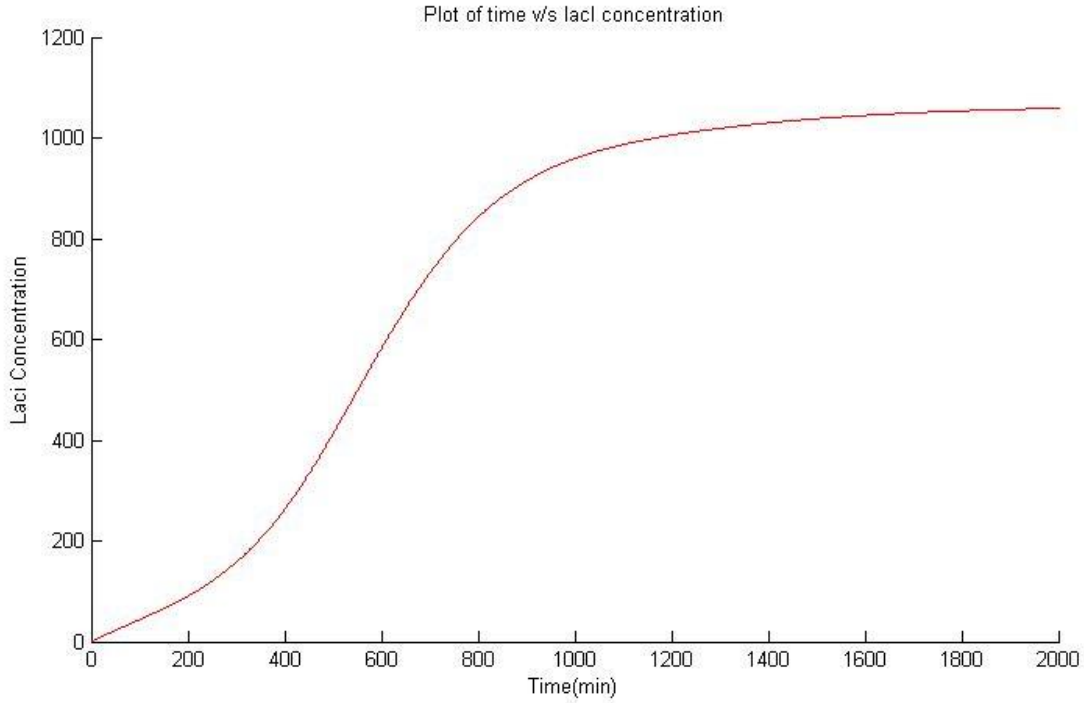
.

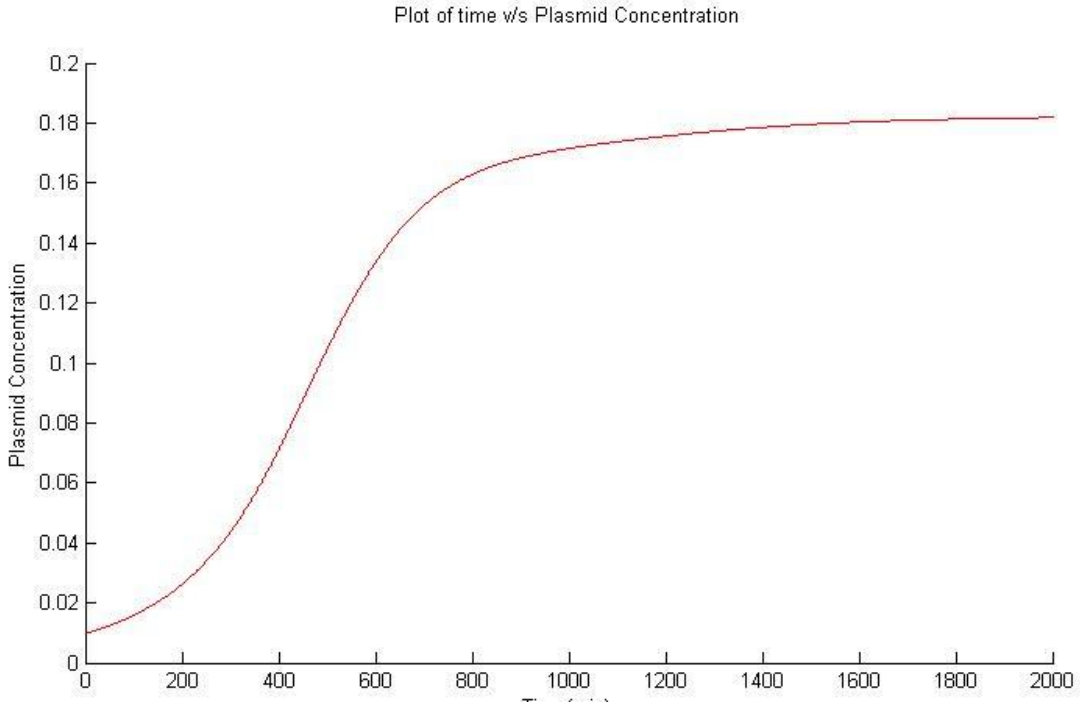
mu 0.496259

Std. Err 0.0138032

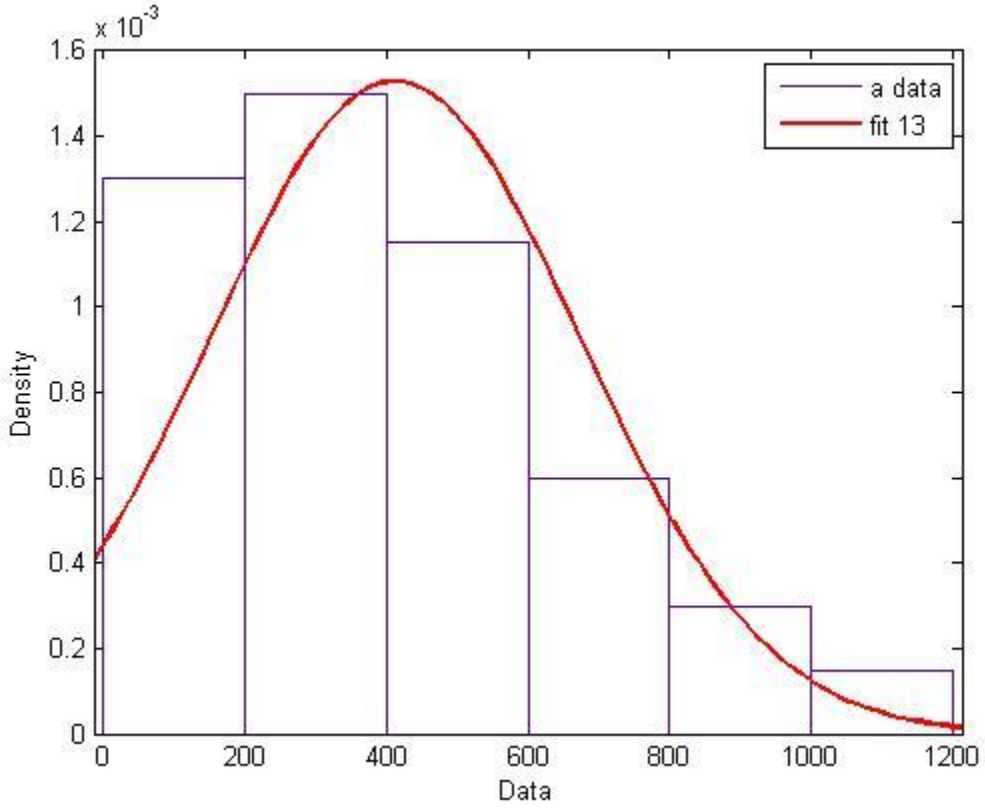
Strain-3(SISO_CN):-

The time response of the strain with copy number regulation is characterized below.





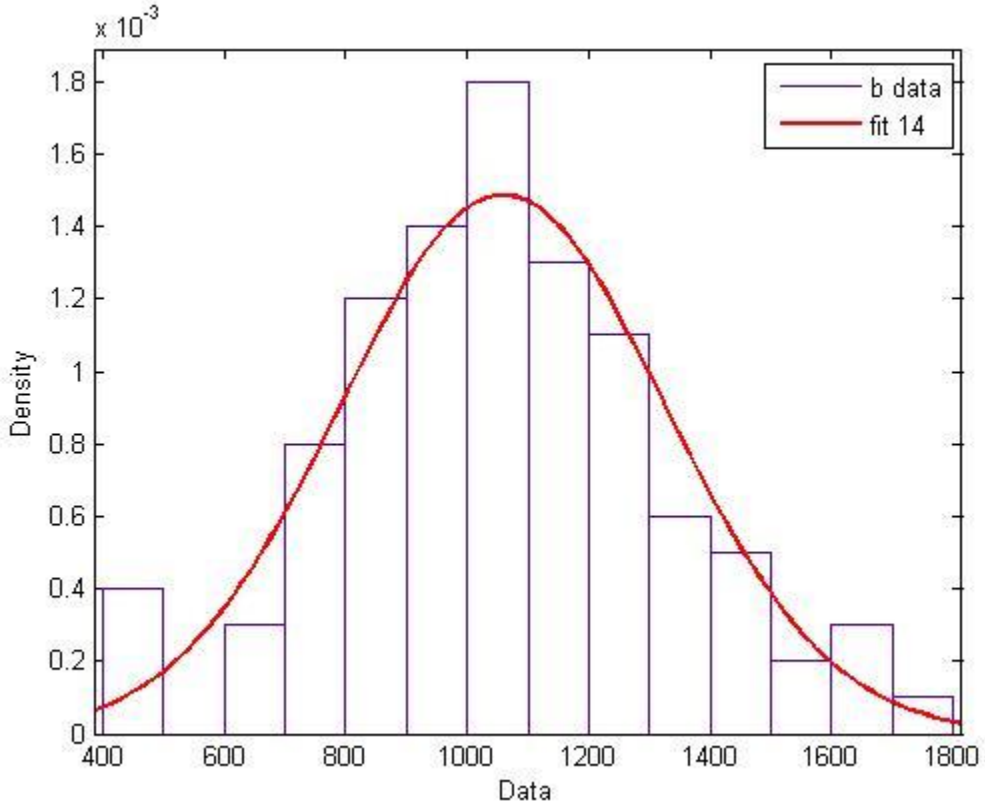
The distribution of lacI concentration at t=500 min obtained by 100 runs



mu 413.063

Std. Err. 26.1133

The distribution of lacI concentration at t=2000 min obtained by 100 runs

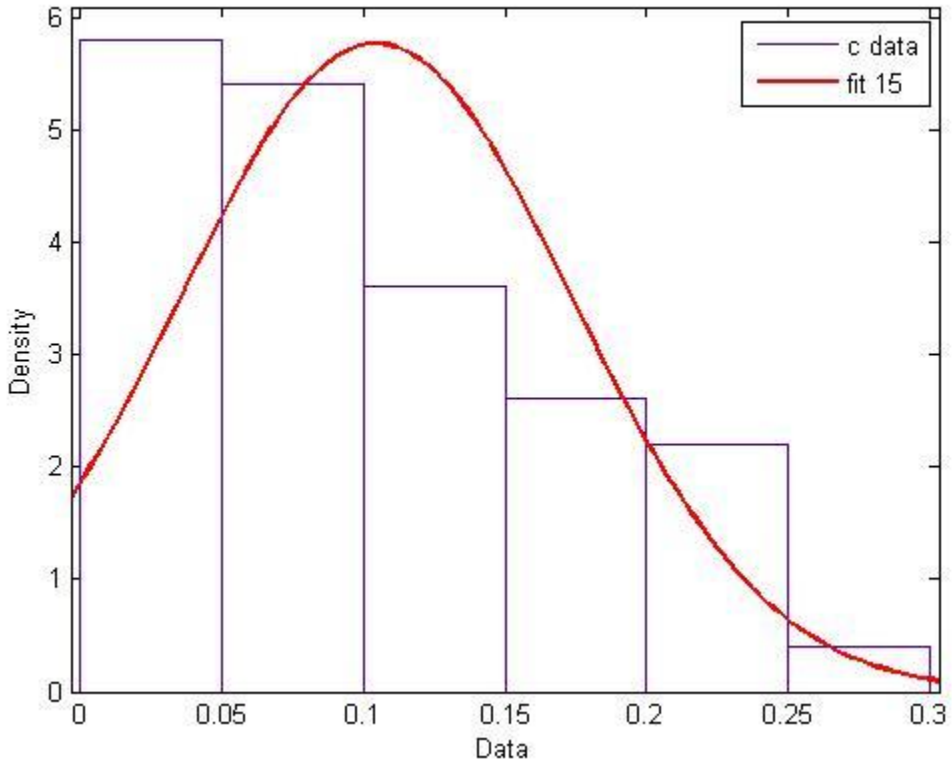


.

mu 1058.75

Std. Err 26.8515

The distribution of plasmid concentration at t=500 min obtained by 100 runs

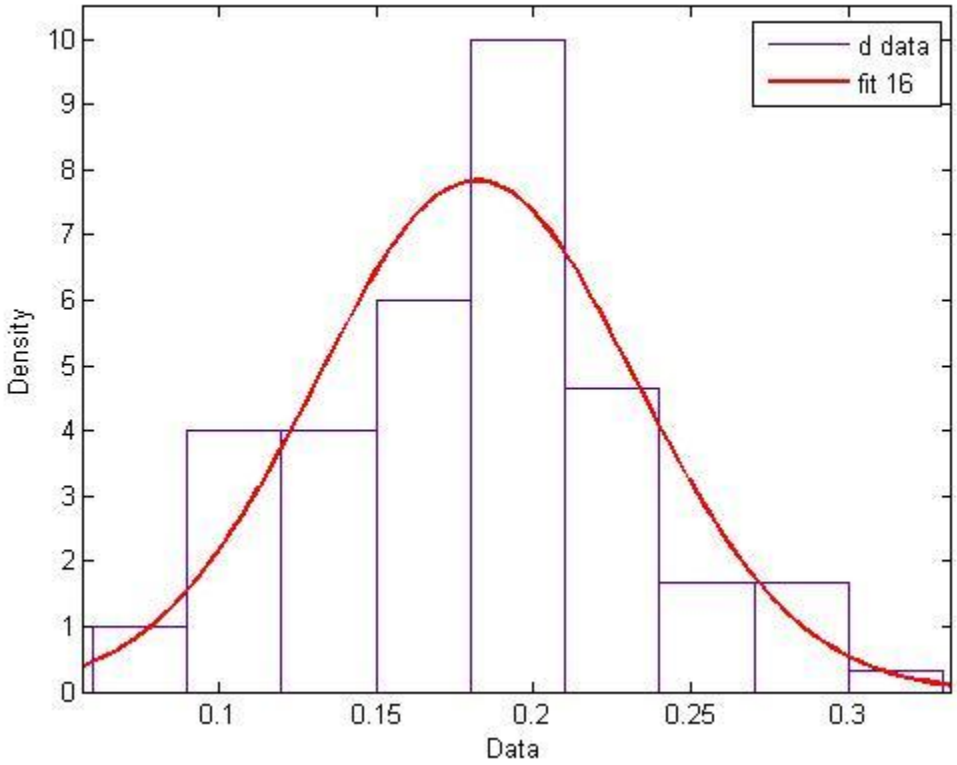


.

mu 0.104611

Std. Err 0.00692011

The distribution of plasmid concentration at t=2000 min obtained by 100 runs



.

mu 0.181828

Std. Err 0.00509243

The means and errors for lacI are tabulated for all the 4 strains are summarized.

Strain	Mean(t=500)	Std Error(t=500)	Mean(t=2000)	Std Error(t=2000)
Open loop	457.865	32.1751	2912.57	95.0793
SISO_lacI	383.837	22.0777	1181.86	23.8863
SISO_CN	413.063	26.1133	1058.75	26.8515
MIMO	360.034	19.4673	872.65	17.5968

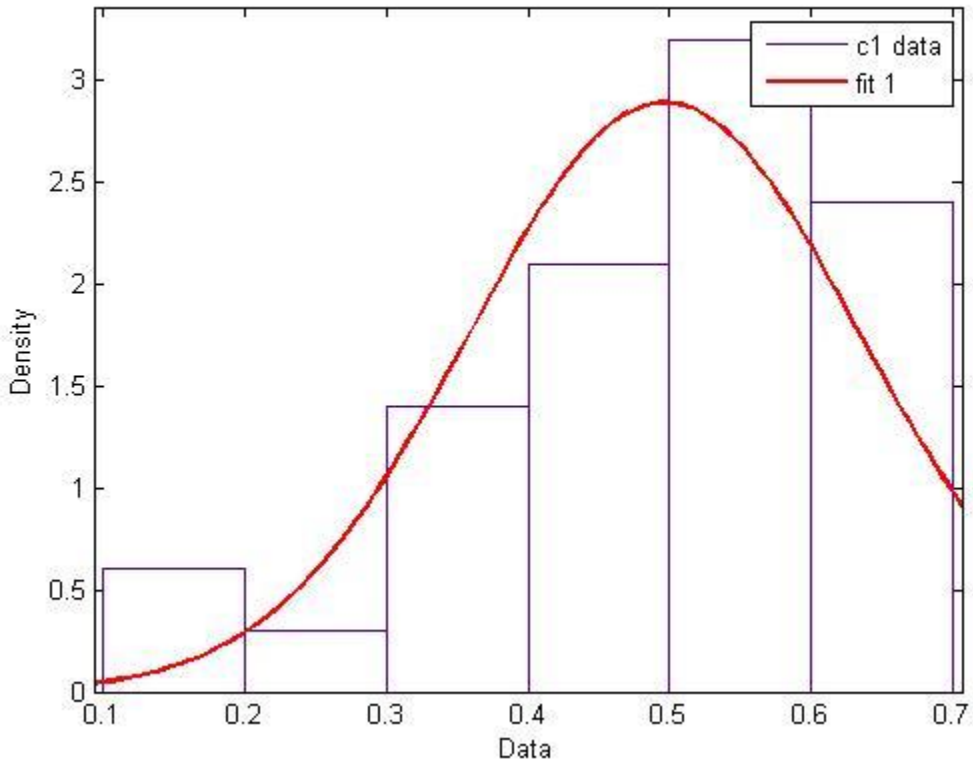
The means and errors for plasmid concentration for all the 4 strains are summarized.

Strain	Mean(t=500)	Std Error(t=500)	Mean(t=2000)	Std Error(t=2000)
Open loop	0.124424	0.0093915	0.496259	0.0138032
SISO_lacI	0.124424	0.0093915	0.496259	0.0138032
SISO_CN	0.104611	0.00692011	0.181828	0.00509243
MIMO	0.107868	0.00739123	0.277137	0.00955624

The distributions obtained for the steady state values of lacI and plasmid concentration at high IPTG are plotted below.

Strain-1 Open loop:

Plasmid Concentration Distribution at steady state for IPTG=100 obtained through 100 runs.

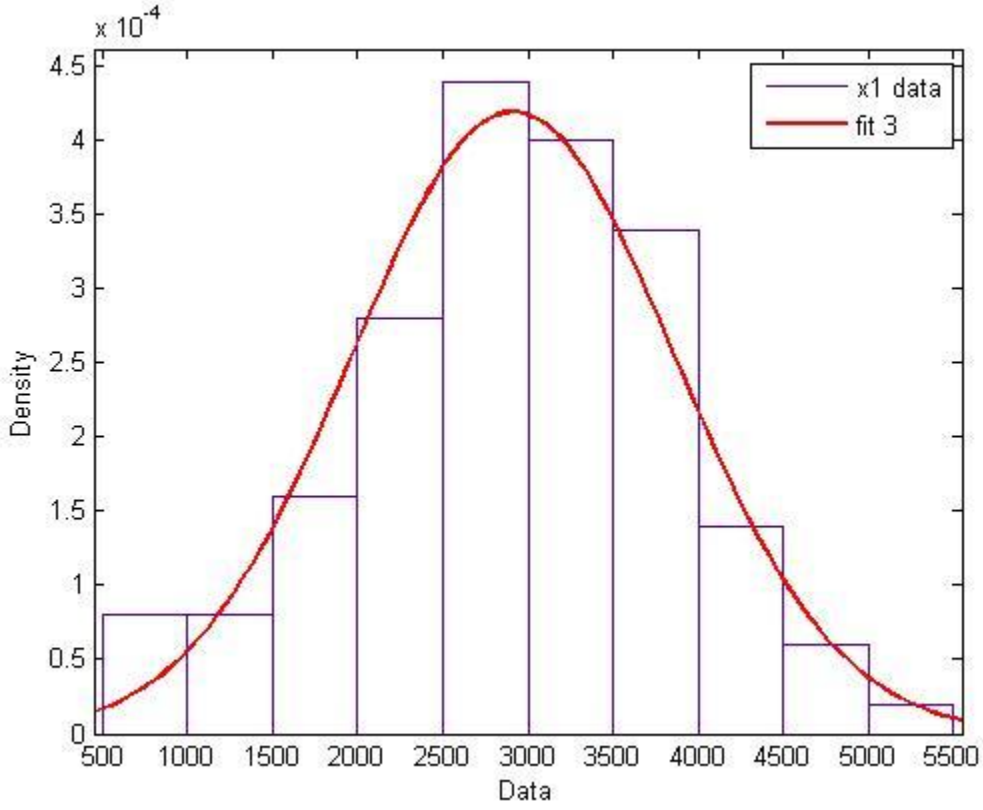


.

mu 0.496259

Std. Err 0.0138032

lacI concentration at steady state for IPTG =100

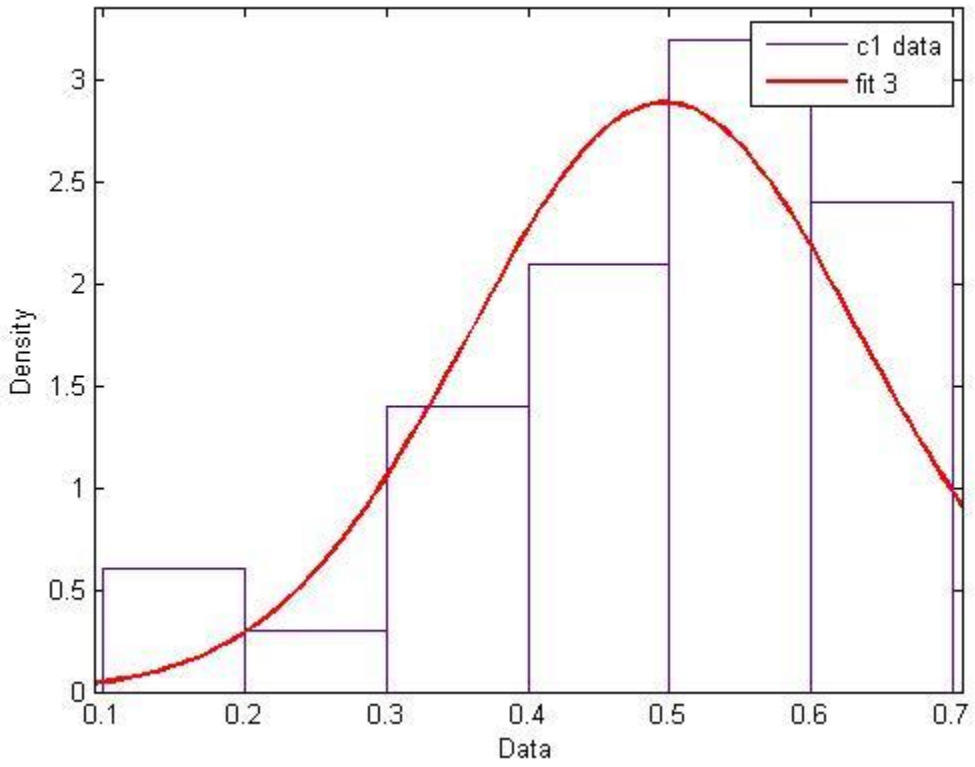


mu 2912.57

Std. Err 95.0793

Strain-2(SISO_lacI):

Plasmid Concentration Distribution at steady state for IPTG=100.

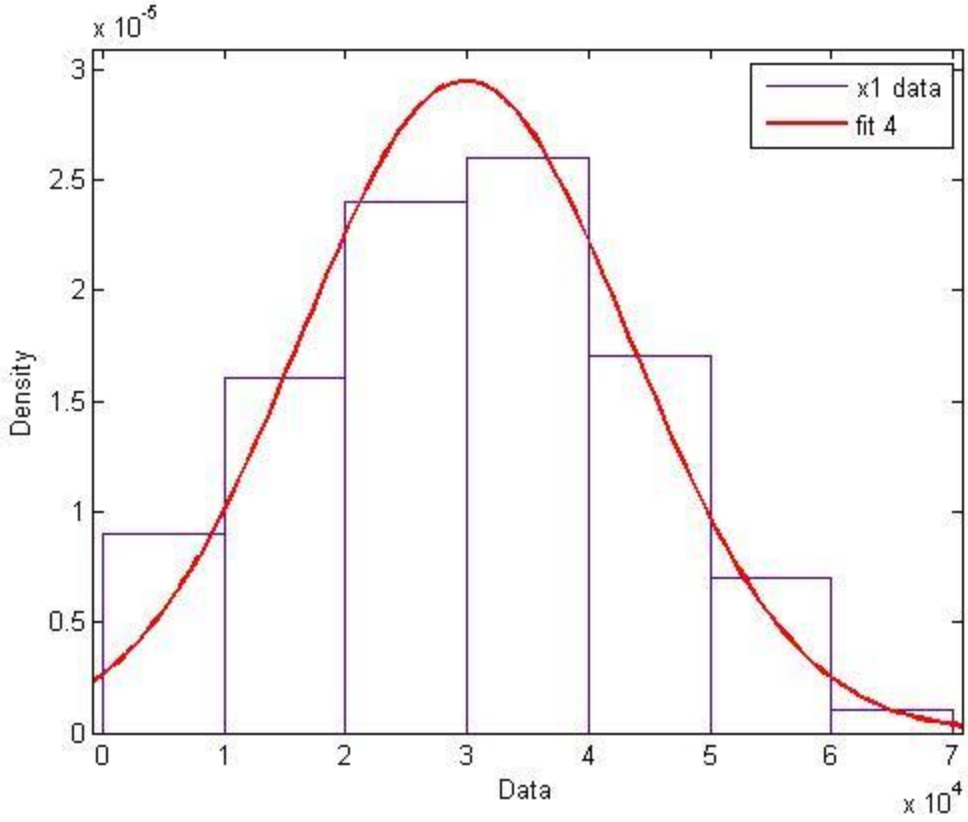


.

mu 0.496259

Std. Err 0.0138032

lacI concentration at steady state for IPTG =100

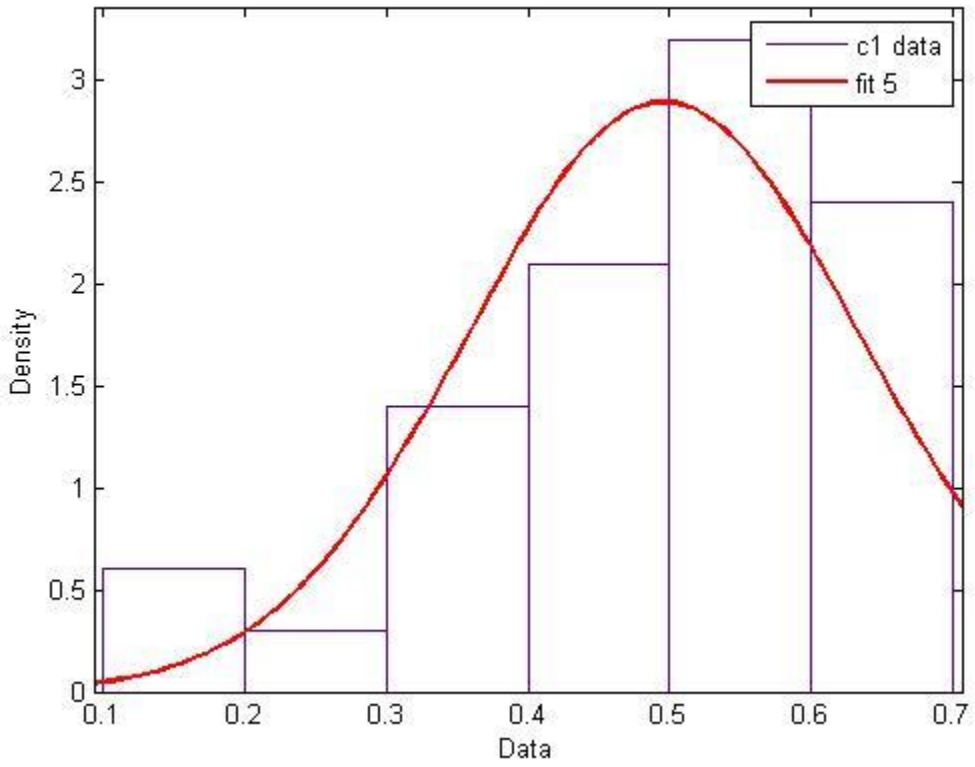


mu 29840.2

Std. Err. 1354.97

Strain-3(SISO_CN):

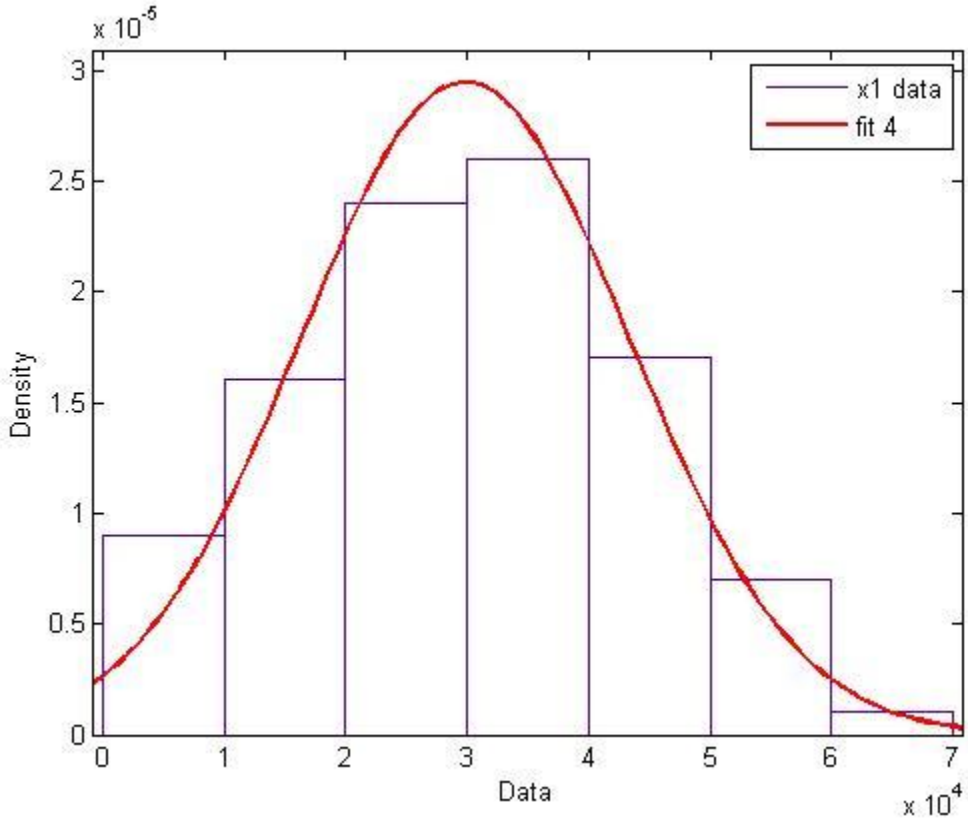
Plasmid Concentration Distribution at steady state for IPTG=100.



mu 0.495936

Std. Err. 0.0137955

lacI concentration at steady state for IPTG =100



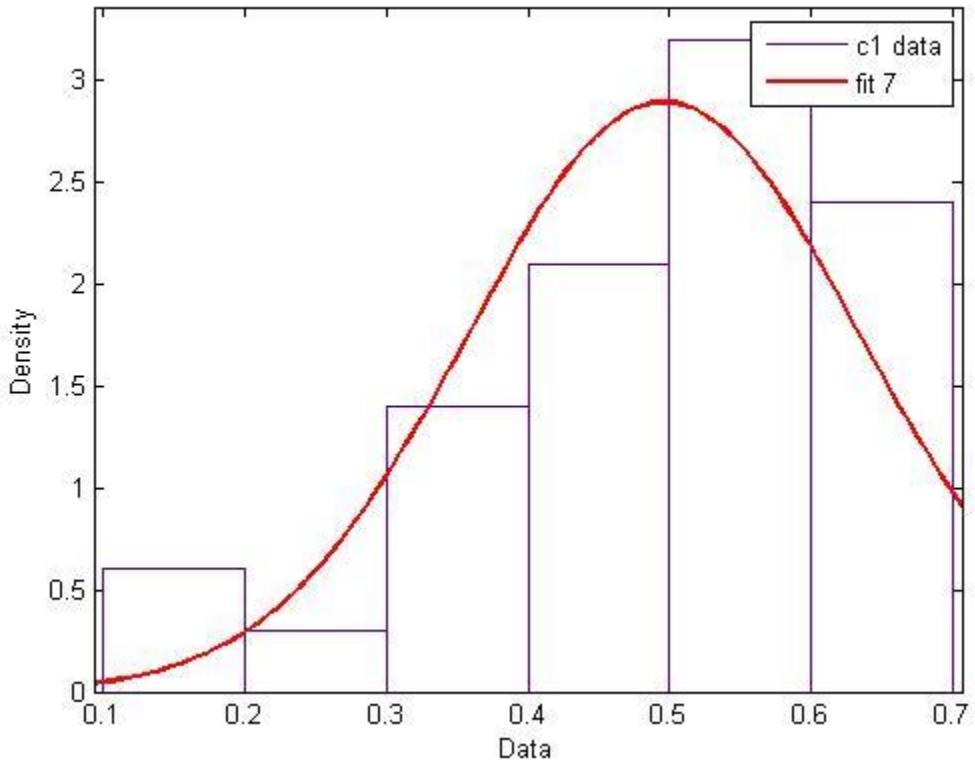
.

mu 29854.7

Std. Err 1356.58

Strain-4 MIMO:

Plasmid Concentration Distribution at steady state for IPTG=100.

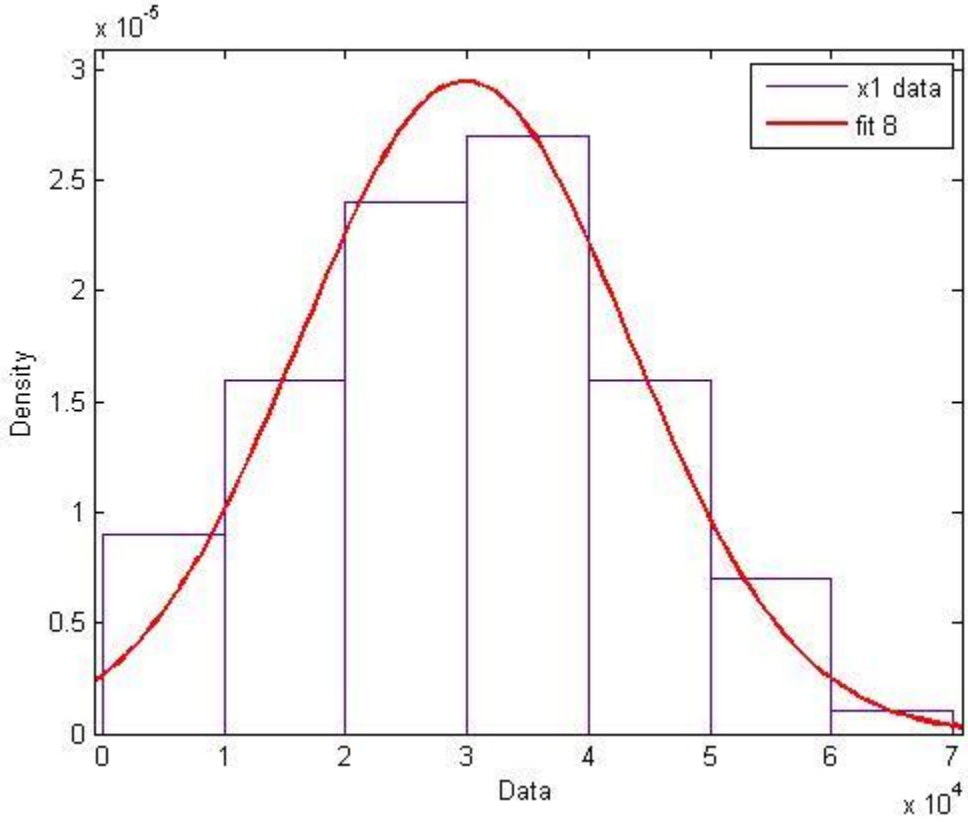


.

mu 0.495937

Std. Err 0.0137955

LacI concentration at steady state for IPTG =100



.

mu 29821.9

Std. Err 1354.07

Steady State LacI Concentration

Strain	Mean(IPTG=0.01)	Std Error(IPTG=0.01)	Mean(IPTG=100)	Std Error(IPTG=100)
Open loop	2912.57	95.0793	2912.57	95.0793
SISO_lacI	1181.86	23.8863	29840.2	1354.97
SISO_CN	1058.75	26.8515	29854.7	1356.58
MIMO	872.65	17.5968	29821.9	1354.07

Steady State Plasmid Concentration

Strain	Mean(IPTG=0.01)	Std Error(IPTG=0.01)	Mean(IPTG=100)	Std Error(IPTG=100)
Open loop	0.496259	0.0138032	0.496259	0.0138032
SISO_lacI	0.496259	0.0138032	0.496259	0.0138032
SISO_CN	0.181828	0.00509243	0.495936	0.0137955
MIMO	0.277137	0.00955624	0.495937	0.0137955