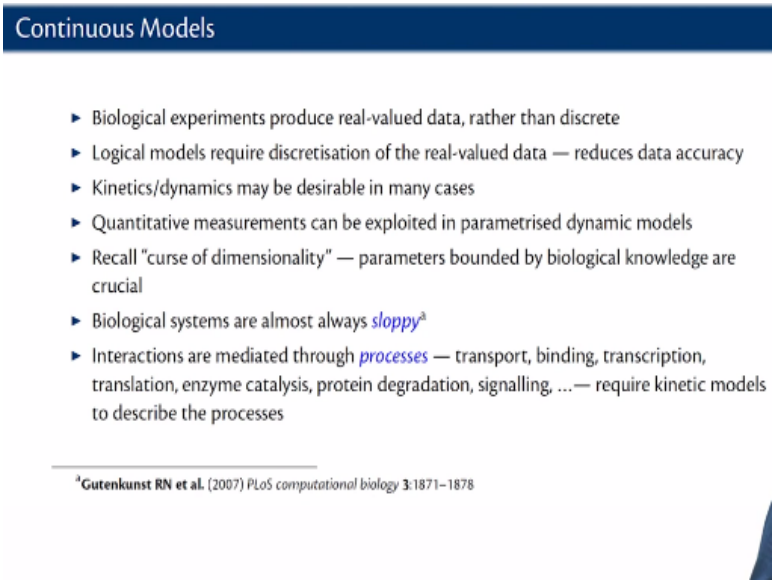
Computational Systems Biology
Karthik Raman
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture – 88
Modelling Gene Regulatory Networks

In this video, we will look at continuous models, these are more similar in spirit to the dynamic models, ODE's models we saw earlier on but focusing on gene regulation, let us quickly look at continuous models, this is something a very familiar with in any case but they are also useful in the case of regulatory network modelling.

(Refer Slide Time: 00:29)



So, we have obviously real value data than discrete from most biological experiments and you have to somehow use a threshold or something like that to discretize whatever real value data you get right, which reduces accuracy that is a problem and in many cases, you may want kinetics, right we discuss this in the very beginning of the course that your choice of modelling tool is going to be dictated by the kind of questions you want to answer.

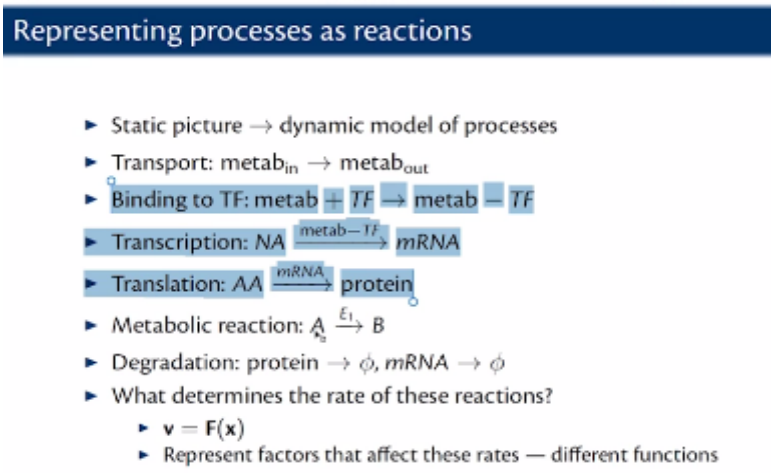
Do you want to do kinetics, do you want to do steady state, they want to just identify who interacts with whom and so on and you know if you have a parameterized dynamic model, you can exploit or infer these quantitative measurements and of course, you have the challenge of the

curse of dimensionality, right, if you have a very large number of parameters you can never fit them, so you have to you know unless you have some biological knowledge?

I know that this parameter is in this range, I know that my KM lies between 0.01 and 0.08, right unless I have that kind of information, it becomes very difficult to estimate the parameters and this systems can also be very sloppy, right, so in the sense that they can show the same behaviour for a very large variations in parameters and so on and what are all the process involved, so you have many complex processes that are involved in gene regulation, right from transport, binding, transcription, translation, catalysis, degradation and also some signalling events and so on.

So, all these need to be described using kinetic models and we are familiar with the same good old kinetic models, what are the classic kinetic models that we need to worry about; Michaelis-Menten kinetics, Hill kinetics, mass action kinetics, invariably you can represent all of these using one of these classic kinetic models.

(Refer Slide Time: 02:28)

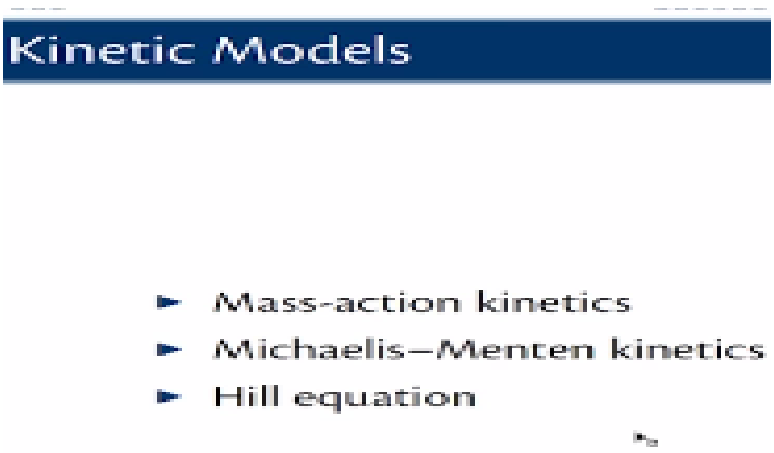


So, you need to basically take a static picture and transform it into a dynamic model of processes, so maybe a transport will look like a reaction of this sort, binding of a metabolite to a transcription factor, so transcription factor will take a nucleic acid and you know transcription involves the conversion of a nucleic acid to an mRNA in the presence of transcription factor, then translation amino acids to protein when mRNA is present.

And then a metabolic reaction that we are very familiar with, in the presence of an enzyme, A will go to B and in degradation, protein degrades mRNA degrades and you can basically write a rate function which captures the rate of all these reactions, right so you need to represent all the factors, they might be allosteric inhibitors, they may be allosteric activators, other effector molecules, ions, you know sensitivity to pH whatever.

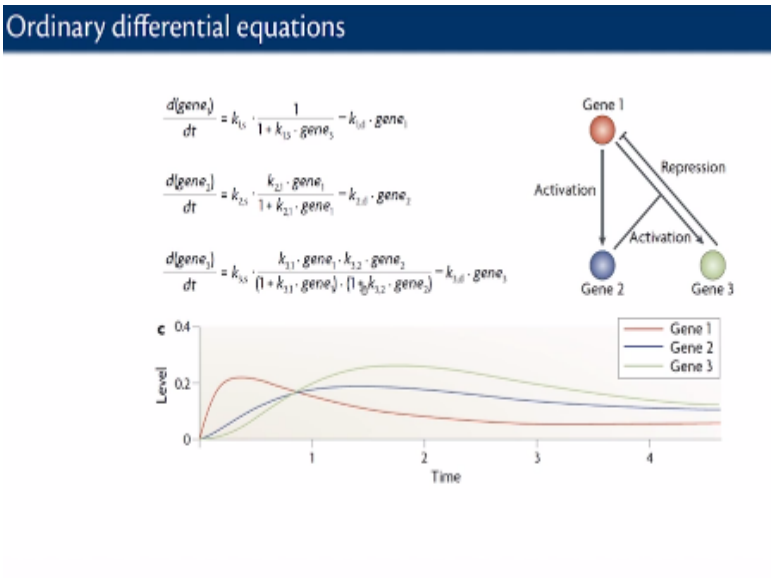
All these things have to be incorporated into your dynamic model one way or the other, so this very same system can be modelled using Boolean models if you want, they can be modelled using continuous models, you could also use FBA occasionally but you know FBA is not good enough to typically model some of these kinds of system even though, you can show them as reactions, you know it is very difficult write mass balance for a protein and so on.

(Refer Slide Time: 03:59)



Because you are talking about several amino acids, polymerisation and things like that and things like. So, the classic kinetic models again are mass action, Michaelis- Menten and Hill.

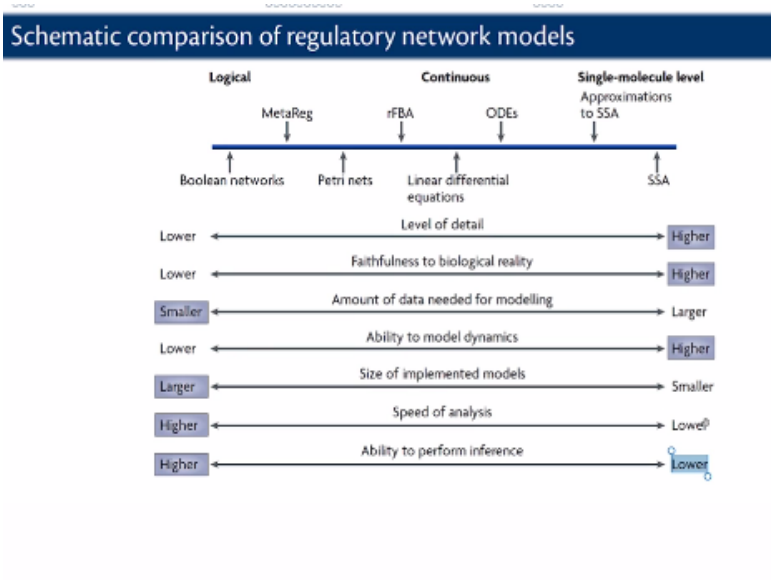
(Refer Slide Time: 04:04)



Let us look at a simple system here, so you have a gene 1 activating gene 2 and gene 1 again activating gene 3 and gene 1 and gene 2 activating gene 3 and gene 3 represent gene 1, right, how do the equations look like, so dgene1/dt is some; so in all these cases, you see that there is a negative contribution from the same gene, so there is a sort of degradation rate that is captured by these constants, right dgene/ dt is includes a minus self-term.

In addition to that, you have K1 * 1+; this is more like your Michaelis- Menten, you can have more complex equations like this and so on and you will end up with time courses that can be captured, the catch of course is I have to estimate 1, 2; how many; 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 different parameters for a simple system with just 3 genes, right it can become very challenging.

(Refer Slide Time: 05:15)



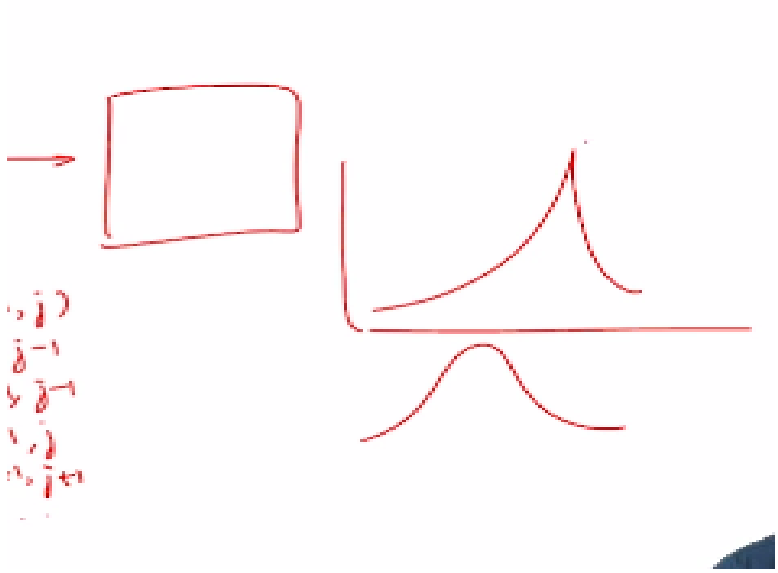
So, to summarise there are several methods that are available and each of them vary in their ability to capture various details. So, for example if you have to look at the level of detail stochastic simulation are you know, very detailed models but the problem is that the models themselves are very small, stochastic; it is something like you know a transcription event is often could be modelled as a stochastic event.

Because you will have like only a few transcription factors in the cell and maybe a few RNA polymerase molecules in the cell right, so you are going to be worried about whether this binds to DNA and so on, so you all going to look at concentrations and so on and we will start looking at distributions of numbers, so how many polymerases are bound by transcription factors verses unbound and so on and so forth.

So, how faithful are they to biological reality; obviously, at the level of stochastic models, they are very faithful to biological reality but you need a very large amount of data for modelling and even though, you have a very high ability to model dynamics, you can only implement very small models, it is clearly see that there is a lot of trade off involved here and this according to me is one of the most important takeaways for you from this course and this series of lectures, right.

When do you use what kind of model, of course, speed of analysis again stochastic models will take a very long time to simulate because you have to simulate various realisations and so on and ability to perform inference, this is probably the most important thing and this is very low in terms of stochastic models because we are talking about system level inference, you can make very accurate predictions about a single small system.

(Refer Slide Time: 07:26)



But if you are looking at a system level inference, the ability to do that really declines with the size of the; with the small of the size of the system, lower is your ability to make these kinds of predictions. So, you will have a distribution, you will know that you know this is going to be my distribution or this is going to be my distribution of whatever molecules and you know number of molecules bound to transcription factors.

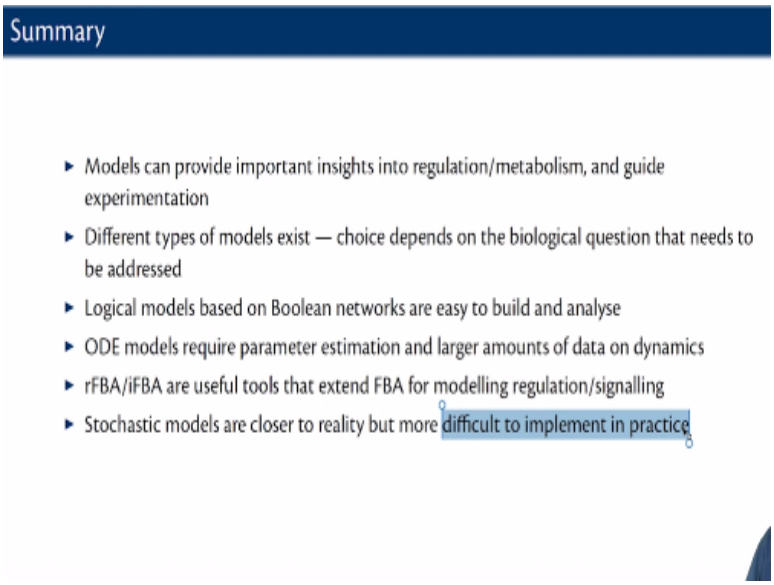
So, what is the most likely state, you can figure that out so, this slide essentially tells you according to several different access, right so, from level of detail, faithfulness to biological reality, amount of data needed, right this is easier to build Boolean networks, we do not need too much data to build them, rFBA and those things are somewhat right in the middle, right you can even put FBA down here, although we are talking regulatory modelling here.

But in general and here in fact if you looked at the entire course, your networks will go somewhere here, the networks you studied in the very first module, right and continuous is of

course here but then, you are sort of limited by the size of models that you can build, models are going to be smallish when you are looking at ODE's, so it is clearly a trade-off and you will have to you know make the sacrifices necessary for the question that you want to answer.

We are more interested in modelling dynamics then you are limited to this half of you know models to this right of this were, right if you are very interested in you know making very large system and prediction, you want to look at somewhere here or if you want to be very accurate, you want to look at somewhere here but then again, you would not be able to simulate a very large system.

(Refer Slide Time: 08:57)



So, you clearly have trade-offs to make and figure out, the models can provide important insights into regulation and metabolism and ideally guide experimentation and obviously we have different types of models and the choice depends on what is the question you want to address, you want to know if this interaction is going to happen or not, you may survey with a very simple model.

Do you want to know what is the rate at which this signalling pathway kicks in when this sort of shock is given to the cell, then you will need a may be a dynamic model right, so depending upon the question that you want to answer, you need to pick the right kind of model and logical

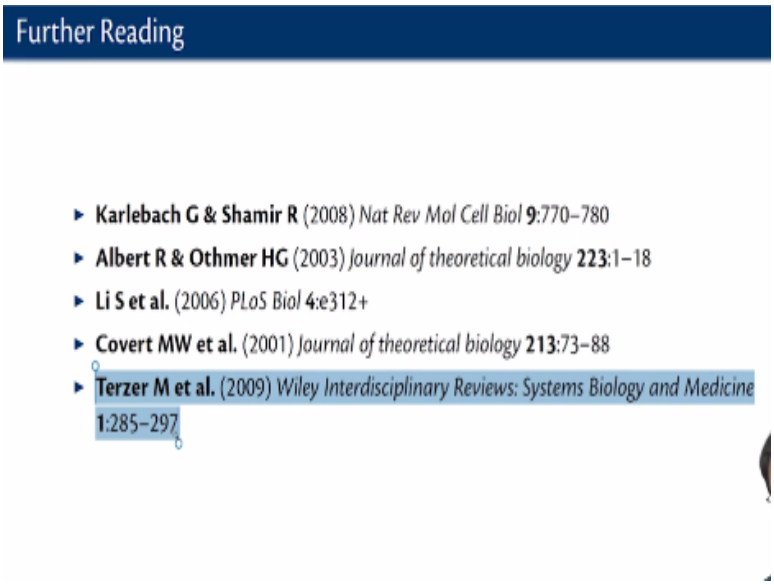
models based on Boolean networks are easy to build and analyse and they are somewhat quite powerful?

We look at some hands-on in the next class and ODE models are great, excellent for doing dynamics as long as you can surmount 2 things; one, availability of data; 2, parameter estimation, right and both are sort of I would say not exactly orthogonal but both can be issues; even if you have a lot of data, if you have too many parameters to estimate, you are not going to go anywhere.

And if you do not have a lot of data, you are already at a very big disadvantage to start off with and tools like rFBA or integrated FBA, there are ways to you know integrate genomic data into models and so on, so all these are useful because they can extend FBA for modelling regulation and even signalling if you want, so essentially, marry FBA with some discrete modelling framework.

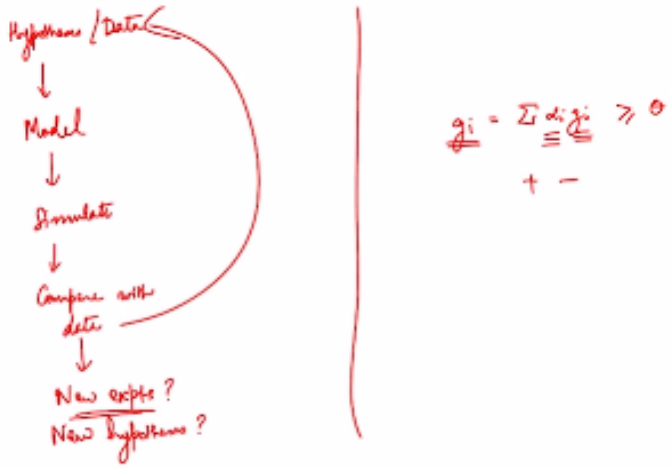
You take FBA and connected to say you know Boolean rules for various regulators and so on, which is what is rFBA, you can do it similarly for signalling and so on. Stochastic models are very useful, I mean very powerful to make realistic modelling and simulation where they are very difficult to implement in practice and will they can make only small; you can make only small size inferences.

(Refer Slide Time: 11:04)



So, there are very nice stochastic models of chemotaxis and things like that but you seldom have stochastic models of an entire cell and so on. So, these are some very good references, I will post them to the class website, so this is a very good review that overviews FBA and a lot of what we have studied in here, this was one of the classic Boolean network modelling papers, these two, this is the Abscisic acid signalling paper, this is a classic rFBA paper and this another nice review.

(Refer Slide Time: 11:42)



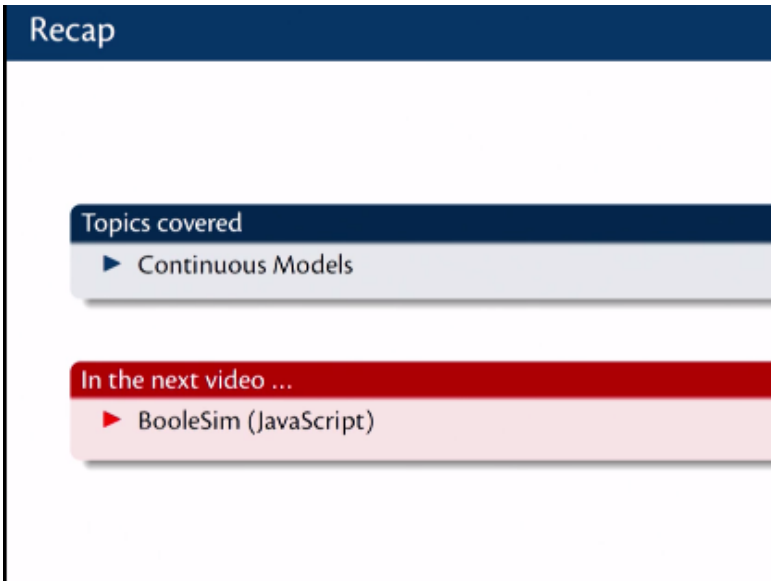
So, you can basically figure out where all the action is, where all your doubts are and then so this is in fact one likes to go about modelling in general, right. So, you first you know have some start with some hypotheses right, or if you have data, then you have a model, you simulate or

validate and compare with data right, does it agree or not, right, so depending on this, you basically figure out what are the changes or you know in fact or can this tell you new experiments or new hypothesis.

This would be the best contribution, right can you perform the most informative experiment, the one useful experiment that you want to do based on the current set of data and observations simulations. **"Professor – student conversation starts"** so, there are other ways to simulate, so you may just you know, so you can just simulate, so let us say gi is some; so you can just say the state of this is greater than some threshold theta, right.

This will be a different way to simulate, it is not a true logical, it is not a Boolean network based model but if you have an adjacency matrix, you can also simulate, so I have like 10 regulatory genes each having some weight, some having + weight and some having - weights and if the total effect is greater than some theta, I will assume that this is turning on or off or something like that. **"Professor – student conversation ends."**

(Refer Slide Time: 13:35)



So, in this video I hope you got a good overview of continuous models and also analyse summary of the different kinds of models that one can use for modelling gene regulation and when one likes to use them and what are the kind of efforts involved in modelling and what the kind of outputs you get from the model and so on. In the next video, we will do a lab where in

we use the simple JavaScript based tool called BooleSim to simulate a Boolean network, a simple Boolean network.