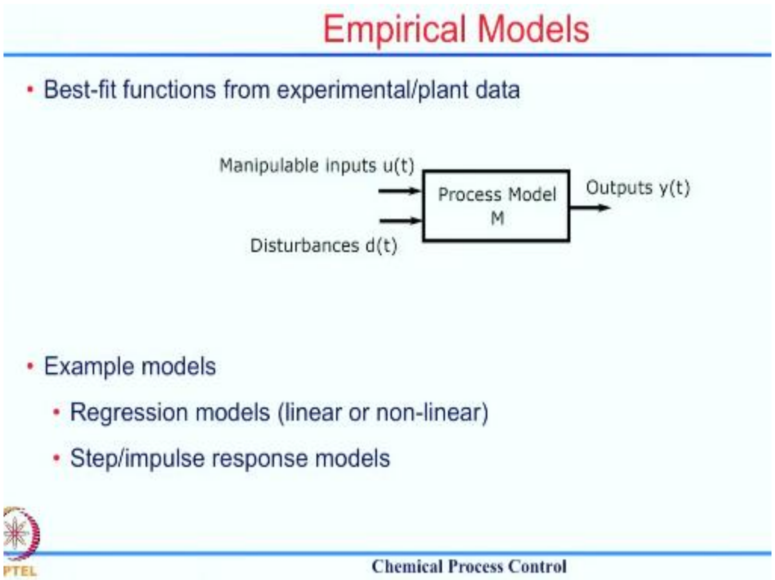
Chemical Process Control

Prof. Sujit S. Jogwar
Department of Chemical Engineering
Indian Institute of Technology-Bombay

Lecture - 07
Empirical and Gray Box Models



Let us now see what are the other options which are available with us in case the first principle dynamic model has limitations for the process under consideration. In that case, we go with what is known as an empirical model. The empirical model one which will be developed with the help of experiments. As we had seen earlier, a dynamic model is that, which gives you a relationship between inputs and output as a function of time.

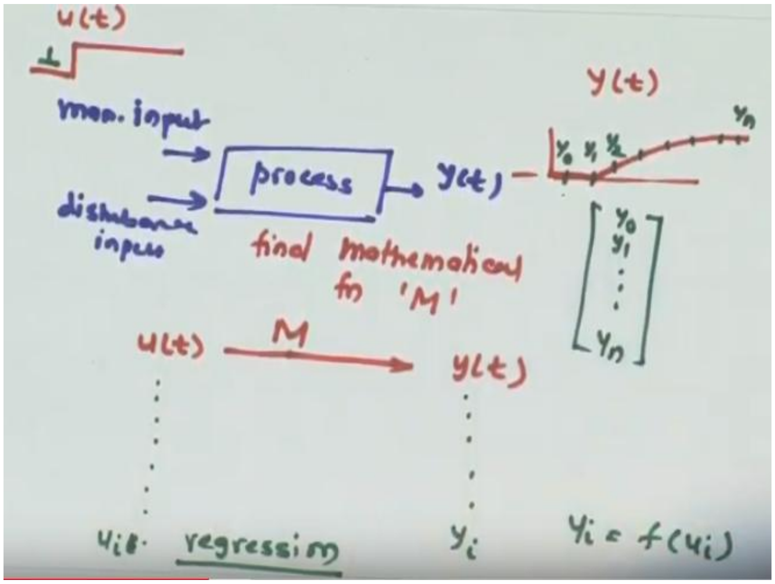
To generate an empirical model, for a process already built, you give changes in the inputs which may be a manipulated input or a disturbance input and then capture the response of the output as a function of time.

(Refer to Slide Time: 01:21)

Let us say we have a process with manipulated input and disturbance input giving an output. We give some change in the input. Let us say we give some step change in the value of the input, and we try to record what is the value of the output. Because of that change, the output did not change for some time and later on started showing some response. So, we have how the input changes as a function of time and we have recorded how the output changes as a function of time and then we tried to find the function which is the process model which will convert this input into the output. So we do not need to necessarily know what is inside the process, what kind of equipment is there, what physical laws it follows. All we are interested in giving some change in the input and then record the output and what mathematical function would give me a relationship which says that if a particular input is given to the process, the output which we saw is same as what will be predicted by using first principle model.

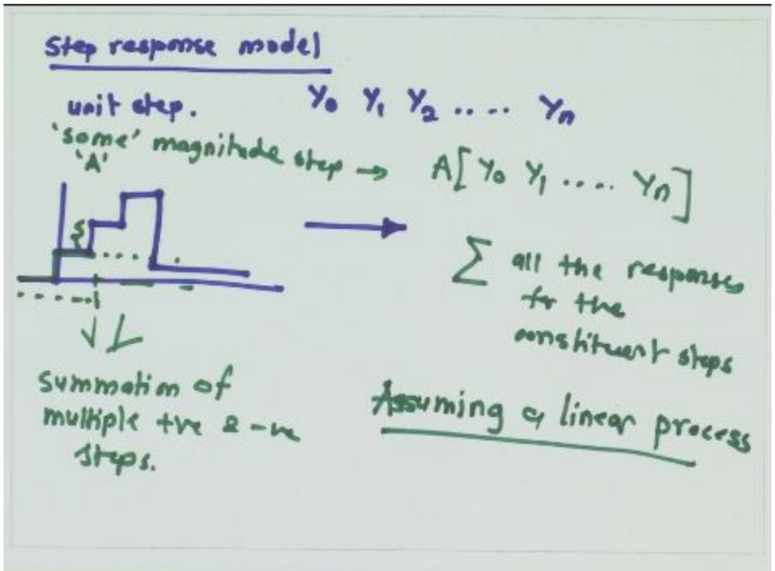
There are multiple ways in which this empirical models can be formulated. One is the simple regression model. We have the input. We have data about how the input changed. These are all the values of u_i, and then we have the values of all *y's* which are y_i, and we can simply do regression which will tell me y_i as a function of u_i. So the simplest way it can be done is a linear regression or a nonlinear regression model.

The other way we can capture the response, is for a step change of a certain magnitude. Let us say this change was of magnitude 1. That is known as the unit step change. You capture the output at different time instances. So let us say these are the different time instances and you note down for a unit step change in the input what are these coefficients. So we will call it y_0, y_1, y_2, all the way up to y_n. So listing these coefficients will tell me how does the system respond for a unit step change. So such kind of a model where we tabulate the values of y_0, y_1 up to y_n , would be known as a step response model.



In the early advanced controllers, these type of models were used to capture the response of a process. So the idea is, if there is some input which is given to the process, you can always write it down as a summation of multiple steps and then assuming the process to be linear we can just add the responses corresponding to each of the inputs. Let me explain this phenomenon again.

(Refer to Slide Time: 05:17)



So step response model says that if I have a unit step change, my output will give me a particular trajectory/response. These will be the changes from the original steady state. Now let us say the input was changed to another form. So we can always decompose a particular input sequence as a series of steps. So let us say the first step is of magnitude *a* for a particular time interval. The second step you can say is 0 for a second period. So by addition of multiple steps, which can be written as a summation of multiple positive and negative steps, we get a particular output which is then analyzed. So if some magnitude step is given, let us say A, the response, will be A times y_0, y_1, y_2, all the way up to y_n.

The response will be the summation of all the responses for the constituent steps. So this all is valid assuming a linear process. If you assume that a process is linear, then any input can be written as a summation of multiple positive and negative steps and then the response can simply be scaled and added together to get the final response.

And then lastly these empirical models can be as complicated as the artificial neural network type model as well. Mostly from control point of view, we would not go to the level of artificial neural networks. But regression models or step response models, those are quite commonly used empirical models for control systems. So now let us look at, what are the advantages and limitations of these empirical models.

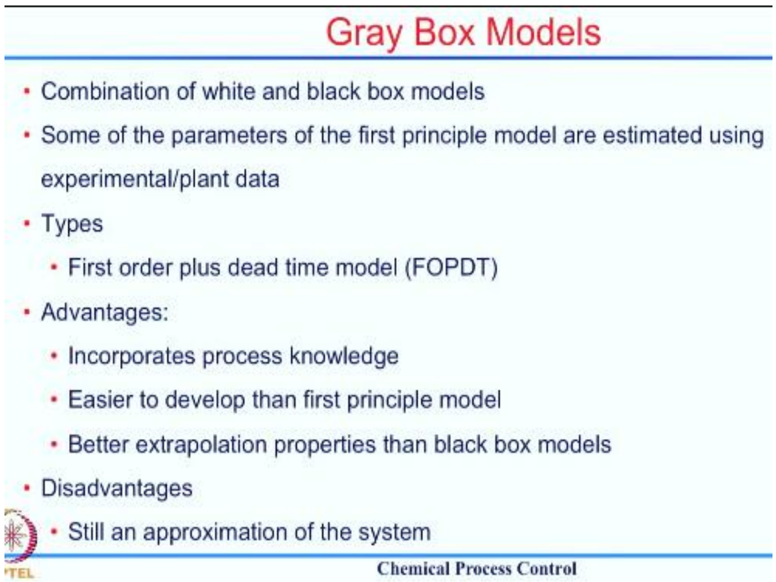
The first advantage is straightforward. These were developed as alternatives to the first principle models. Especially for the cases where the first principle dynamic models are difficult to arrive at these first principle models will be easier to develop. Because in a real plant there maybe hundreds of process variables but out of those maybe few let us say 5 or 6 are inputs and correspondingly 3, 4 are output. So, in that case, we just need to develop individual pairwise relationships between inputs and outputs which may be significantly lower than the total number of equations which you need to write for a first principle model.

Now let us look at what are the limitations, and the major limitation is that these models heavily depend on how these models were arrived at or what kind of changes which were given to the inputs. This is known as excitation of the process. In typical control language, how much did we excite the process, i.e. how much was the step change given to the input and accordingly the validity of the model is more or less of the same order as the excitation. While developing the empirical model, if the changes in the input were not significant, then when you use the model for prediction that time the predictive capability is also limited. So these do not extrapolate very well.

But, depending on what is the main requirement, is it just getting the relationship between input and output which captures most of the expected changes in the inputs, then empirical model may be better than theoretical model especially if the theoretical model is very difficult to get at or the parameters of that model are difficult to get by.

Last type of dynamic model is sort of a middle ground between the empirical model which are known as black box model and the theoretical models which are known as white box model. These are called as gray box models. Gray box models are actually a combination of white and black box models. So you use parts of both these methodologies.

(Refer to Slide Time: 11:22)



We use some parameters from the first principle model. Typically it is the form of the input-output relationship which we get from the fundamental or first principle modeling, and while getting the parameters of the system, we use experiments or empirical data. So in a way it is a combination of the two methodologies and a very commonly used gray box models from a control point of view is known as the first order plus dead time process.

First order plus dead time process.

So it is a "first order + dead time" model. So here, what we do is we assume a form of a function which is,

$$\frac{K_p * e^{-t_d * s}}{\tau * s + 1}$$

So this model form is assumed and when we get the data between input and output we use this data to find the three model parameters which are K_p, tau, and t_d. If it was an empirical model, I would had taken all this data and simply fitted either a regression model or step response model between these two without considering what sort of form that equation would take.

In a gray box model, we assume a certain form as well, and this has the above form and it also has physical significance. So it is not a completely mathematical function which captures the relationship between input and output. There is physical insight into developing that sort of model, and then the data is used only to get the parameters of the model.

In a way it captures advantages of both the theoretical white box model as well as the black box model in the sense that it incorporates processed knowledge, and therefore it kind of gives you better extrapolation capabilities, and it is also easier to develop than first principle model because here you are using experimental data to arrive at the parameters which are very difficult to get theoretically.

If you want to look at the limitation, it is still an approximation. We are assuming a certain form which may not capture the exact reality of the process, and also we are using mathematical tools to arrive at the best values of these parameters which may not exactly match the data, but within a certain acceptable range, it will give you a fit between the input and output. So they are still approximation models, and they will have certain validity range across which their predictive capabilities are valid.

If you want to compare it with the black box model, the gray box model, in that case, would have much larger or wider applicability compared to the empirical model. So in a way, it sits nicely as a tradeoff between the accuracy of the model and the ease of developing the model and that way many a times gray box model is preferred over either a theoretical model or a black box model. So we will take a short break here. Thank you.