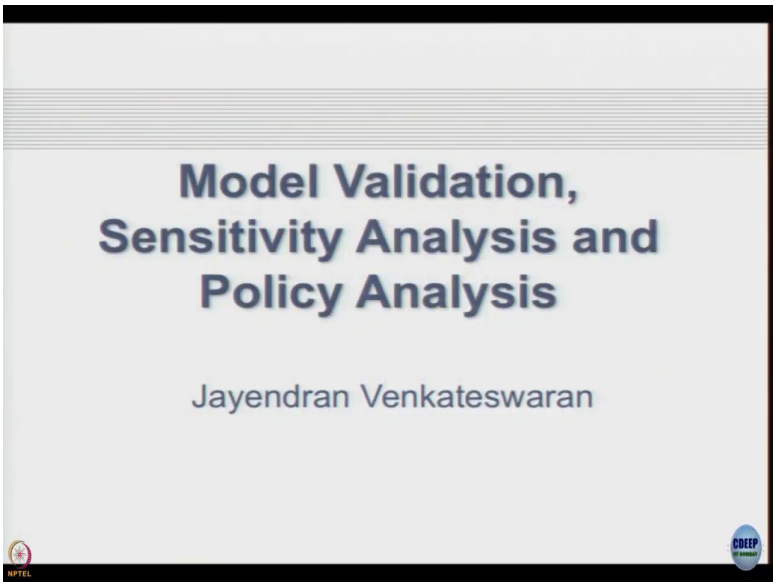
Introduction to System Dynamics Modeling
Prof. Jayendran Venkateswaran
Department of Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay

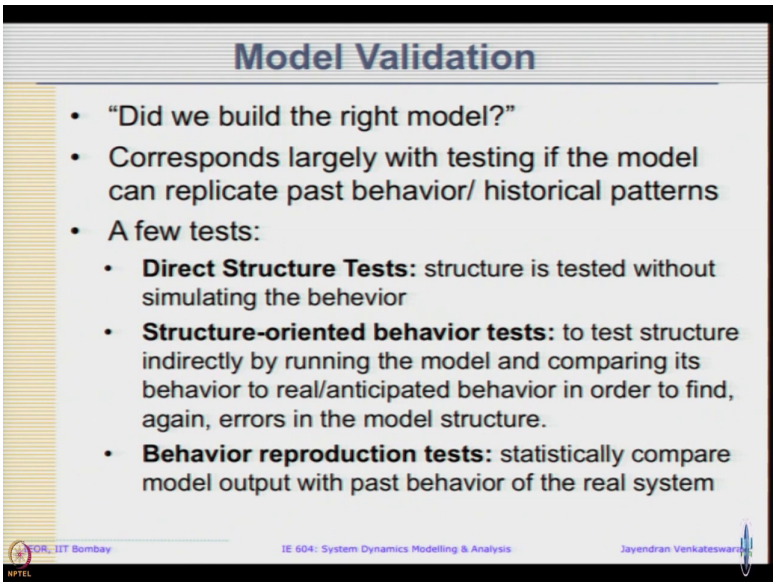
Lecture - 24.1
Policy Analysis
Model Validation and Policy design

(Refer Slide Time: 00:20)



Let us look at more on a model testing today's where we will be looking at topics on model validation sensitivity analysis and policy analysis. I will give a brief on all these topics and like yesterday we will try to look at a couple of examples to reinforce our ideas on what is sensitive analysis and what is this policy analysis. Some of these things you have already been doing in the past where this is giving it a formal name and structure right now.

(Refer Slide Time: 00:45)



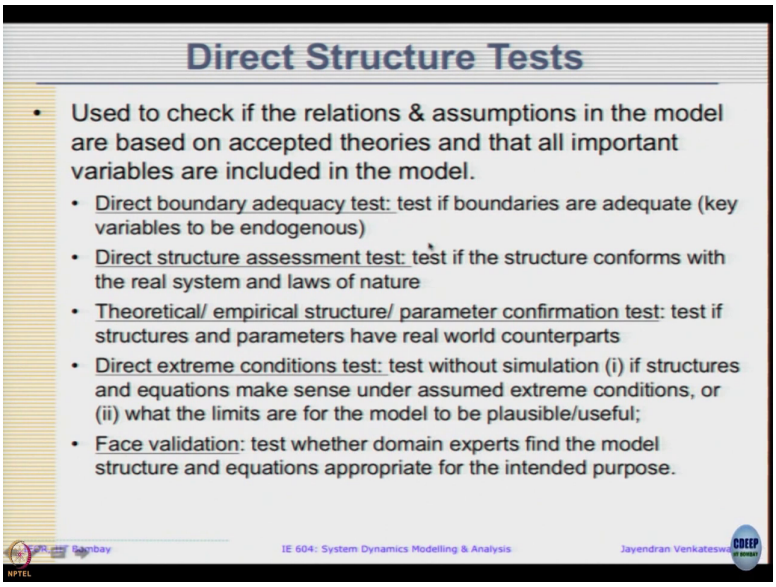
So, model validation talks about whether we build the right model. In verification did we build the model right that was the question; that means, given this specification we want to check whether we captured all those requirements within our model. In model validation we want to ensure that we modelled correct problem. This largely correspond testing of the model can replicate the past behavior as well as historical pattern.

Three broad categories of tests are typically done when we when we talk about model validation. Some of these things may have overlap with model verification. So, that is why when we actually talk about we say model verification and validation has a combined activity with some certain differences are there. So, model validation now first tests when we can try something called as direct structure tests.

In this we are just going to observe the model structure that we have built without simulating the behavior and see whether the structure make sense. Second one is to look at structure oriented behavior tests with our understanding on the structure if we can make some basic simulations that is the behavior mimics the structure that we expect. Suppose we have a positive feedback loop only then we can expect exponential growth is that really happening. Or if you have both positive and negative feedback loops in the system then what kind of behaviors can occur and testing that out on the structure under behavior tests.

Then behavior reproduction test that is to statistically compare the model output with past behaviour. And see whether we can actually represent the past as accurately as possible based on which we can then be more confident that it can affect the that it can be used to predict or analyze what can happen in future. Let us take a little more detail look at this each of these broad tests categories.

(Refer Slide Time: 02:45)



The direct structure test is used to check to the relation assumptions are based on accepted theories and that all important variables are included in the model. Like for example, if you are setting up initial values. As initial values follow with say for example, what is actually being represented as initially available materials and stocks.

If you are setting up some desired variables and in reality people are following Selto's law to use compute those variables there is a module also include that. So, these can be tested without even simulating the model. Was it direct boundary adequacy test? Test if boundaries are adequate that is all the important variables are endogenous. So, if your system has no controlling parameters nothing for you to control then you have attribute all the problems is something external on which you have no control off.

That means, you cannot improve your system unless externally somehow that system changes for the better perfect, so that cannot be the reason right. Even in case of demand in a common inventory system or supply chain demand is considered external. But still people try to influence demand by offering reduction in the sale prices, but timely locating the material at the right locations or opening up multiple sales channels etcetera. So, they do try to influence the demand some how.

So, if that is essential part of doing the business then that has to be part of the entire model itself that is all the first one says. So, as soon as you look at the variables we can check whether should they be exogenous endogenous. If so then how do we include it? Direct structure assessment test, test if the structure conforms with the real system and laws of nature if any like. For example, when deaths occur population has to reduce births occur population has to increase. So, those kind of laws of nature has to be abided by if births fall to 0 then population should not increase.

It is good to check all those what can I say relation within the model that is what we made direct structure assessment tests. Third is theoretical or empirical structure or parameter confirmation test. The test if the structures and parameters have some real world counterparts if we had build the new classical model on stock management structure.

If you recall we had two stocks in that one is a work in process or supply line then inventory. And then we took a decision saying that inventory is going to be adjusted based on desired inventory and actual inventory. And work in process will be adjusted by desired inventory or desired supply line and actual supply line then we also had a forecasting component.

So, it is good to check whether the company is actually doing all that are they company having a forecasting module in it or not. Are they actually considering the inventory information supply information to actually make the decision? So, this is what we mean by theoretical parameter confirmations. Like parameter we had used in that model the parameter called as inventory coverage.

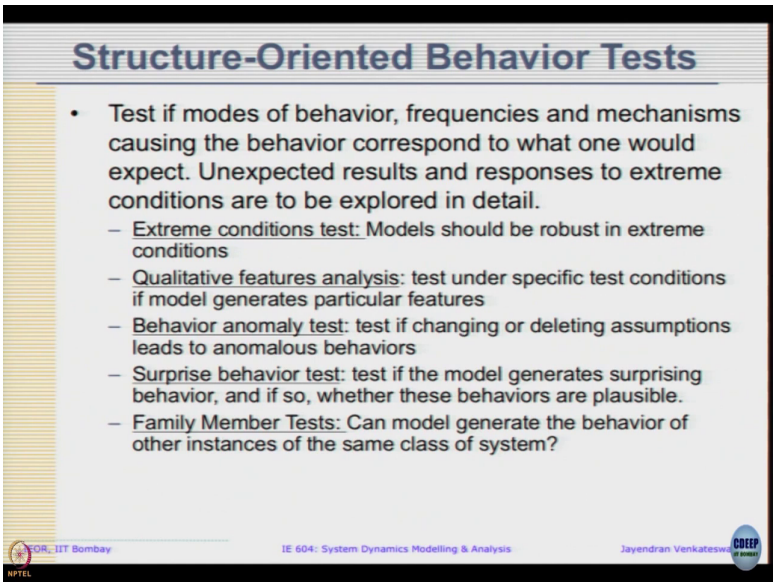
By that we meant how much weeks of inventory we need to have. So, that values you are using here is a realistic to what people are using and be the for simulation purpose you might have used some value. But for when you go to analysis and interpreting the results is there reasonable amount of inventory it hold, so that is what we meant here. Direct extreme condition test, test without simulation structure and equation make sense under assumed extreme conditions. Or what the elements are for the model to be plausible or useful demand is too high or demand falls to 0 the birth rate falls to 0.

So, those kinds of things what can happen or when backlog becomes too high then what kind of things could this model result in can be actually traced back. Face validation, test whether domain experts find the model structure and equations appropriate for the intended purposes. So, now, we are no more in the model verification, when you look at verifications it is in our hands we look at the model we play with the computers etcetera.

Once it comes to validation then we are looking at whether the model is going to be used by someone else out there where we are going to use this model to make the key decisions are we able to convince them. Even if we are consultants we need to ensure that the client actually agrees to whatever the model right. They can say no this does not work this does not (Refer Time: 07:23) does not represent it.

So, that has to be minimize that only come through dialogue. So, that is that part of test is called as face validation when you look it across stable and try to convince them that this model structure actually captures what we intend to it was some broad tests for direct structure tests.

(Refer Slide Time: 07:44)



Second class is called as structure oriented behavior tests. Test if modes of behavior or frequencies or mechanisms causing the behavior correspond to what would one expect right. If you have infectious diseases kind of models then we do expect a shed growth where it is a new product introductions we ended up using a shed growth because of the positive feedback. But, at later time when negative feedbacks has to dominate which results in a shade pattern.

So, the model should actually result in those kind of expected patterns of behaviour. Like if that is large delays in the system then we can expect some oscillations as in indeed occur. If there is no delays still some oscillations that are occurring then we have to figure out what is happening. There is it because of some non-linearity to be introduced which in advertency is causing some unexplained behaviours.

One is extreme condition test similar to what we observed the structure previous step we can actually login on it and see whether the model is actually robust in those extreme conditions or does it throw errors. Qualitative features analysis: testing a specific condition model generates a particular behavior. So, model we try to make it as generic as possible, but in reality we will find that only one scenario is occurred right.

So, under that scenario we know what kind of behavior can occur. So, if you are able to plug that scenario into the model then we get a similar behavior under those conditions. So, what is a behavior anomaly test, test if changing or deleting any assumptions lead to anomalous behaviors is that assumption.

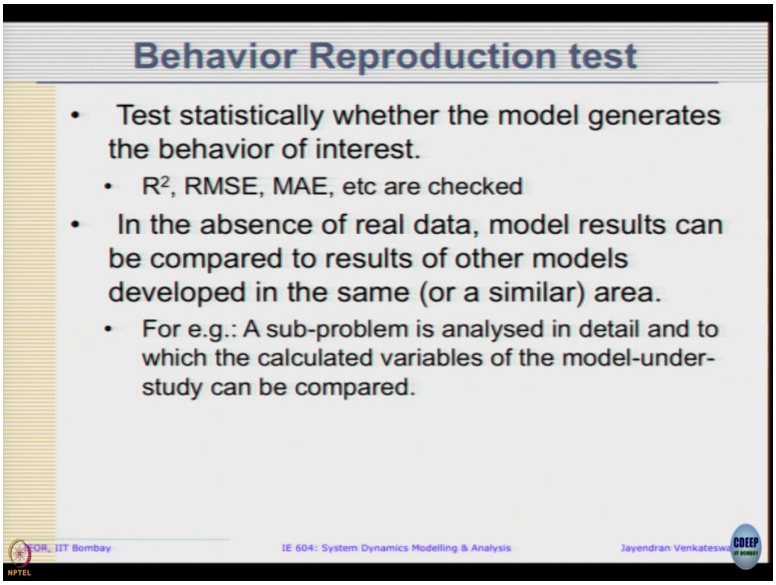
So, critical that it is just made a positive feedback system negative feedback or things like that and does it make sense. Surprise behavior test: test if model is generating surprising behaviors. This is little more difficult to generate unless you generate large amount of scenarios its very difficult to figure out if any surprises occurs. But one is to see whether once you plug in numbers once you look at the results you need to spend time analyzing it that is what today's lecture is about.

Building the model is easy, but that alone is not the end to it. We need to spend some time outside the model to see whether the results actually make sense there is a small blip in the model. You know when we have to zoom in it and understand why that blip is even occurring. Overall the behavior we will be like you are getting a positive feedback. But, maybe it goes up and then there is a small blip and then again it increases then we may round the zoom in to figure out why or what caused that see it is because of moral structure integration issues etcetera etcetera.

Family member tests can model generate the behavior of other instances same class of systems. For example, we are looking at say infectious diseases model. There are so many infectious diseases out there and the model be used for different kinds of infectious diseases. And the model will be used for you said different locations by changing the parameters it still makes sense.

So, more tests variable to do and show that it covers a wider range then more generic the model becomes and more confident that it can be about the model that it can replicate. So, many scenarios and more say more confidence than we can be more sure that what the model is saying we can actually use it for further analysis and things like that.

(Refer Slide Time: 11:04)



Behavior reproduction test, basically this is known as the straight forward ones where this device are more quantitative link line. This is what may appeal to you to test statistically whether the model generates the behavior of interest. Reduce it and reverse when we try to do the modeling for I think it was part of new product introduction.

When you tried to fit the curve with the existing data to see whether you are getting the to compute a values of I think coefficient of imitation and innovation p and q instead of bass

model and then we used it to calculate p and q. But, otherwise suppose we get a some values of p and q we can use it to figure out whether it is able to correctly capture the behavior.

So, this is pretty much done to see that you know once a once a model is built we want to have all the feedbacks. So, that the same can be used to explain what happened in the say the last 50 years or last 30 years. So, if you are able to reproduce that behavior as accurately as possible then we can have concerned that if we can simulate for further 30 years.

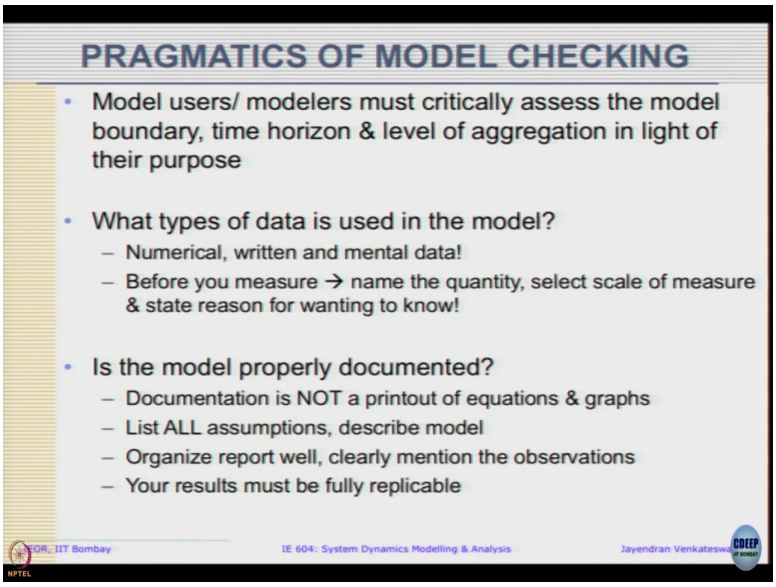
And I am going to use a similar structure then what our model predicts could be reasonable. And if I am going to introduce some changes in the model then I can expect that all the changes are caused because of the intervention we are doing nothing else. So, this another poplar mode of behavior reproduction test is used.

One of the dangers I need to say here about this behavior reproduction is its quite a few it is becomes very tempting as well as a difficult. In the sense we ended up over fitting the with the past data, many times happens that you get the model and tune it so much that it very very accurately predicts the past data. But, in the process the meaning of all those parameters is went for the task.

Like maybe the best fit gave a value of p and q which for the or let us forget p and q let us involve say infectious diseases model. We are looking at contact rate as well as the infectivity parameters can be computed using a past data. Because of us trying to fit the data so hard to that contact rates and infectivity parameters and we got some values.

You may later find that those values cannot exist in nature, like the biologist does not agree that this is the true infectivity there can actually occur, there is no other evidence to it. The problem is because you are over affected to the given data maybe there is other feedbacks that is occurring which you forgot the catch. So, those are small dangers and trying to fit the model so tightly with the existing behavior right, so that also has to be kept in mind.

(Refer Slide Time: 14:07)



Some practical things about checking this models is unfortunately or fortunately only the users or the modelers must critically assess the model boundary. You have to decide what the model boundary has to be how my how long time horizon is to be what is the simulation time step, what is the level of aggregation we need to use. We want to build a factory model or each assembly line in the factory has to be modelled or each machine level we have to model you only need to take pick up the level of aggregation.

Various types of datas using the model you must have seen by now. We have numerical data which is what we are using, but there is written and mental data that we use and try to capture. This comes in the form of relations or multiplications that we do as a fudge factors and gives some values between 0 to 1 these fetch the table function and things like that.

And these are all useful because most of rather everything that we have in the simulation model we would like to have it real world counterpart. So, that we can go ahead and measure it whether it is actually indeed happening or not and use it for further improving of model. So; that means, whenever you use something that we need to have a name for that quantity we need to select the scale of measures that is a units.

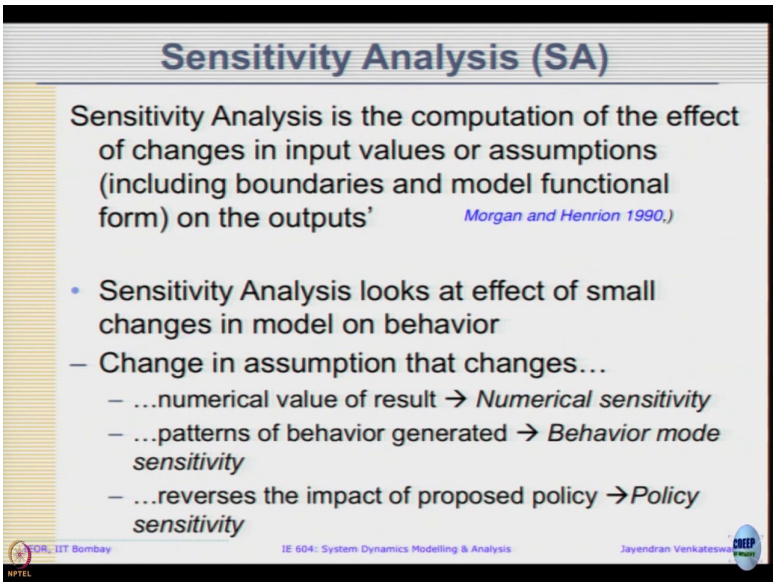
When we state the reason for actually even having those values and that gets closely tied up with model documentation. It is not a print out of equation graphs we need to list the assumptions that is happening like how often (Refer Time: 15:29) is being made. Whether indeed following exponential smoothing based how can I say forecasting method right.

In class we cover exponential smoothing, but the actual while the company may use some other forecasting method. Just because we know only exponential smoothing does not mean we only use exponential smoothing we then try to use forecasting model which mimics what they are having. So, these things come out when you actually write the assumptions to assume that the forecasting exponential smoothing method with, so much parameters and stuffs like that.

These are all very easy to say quite difficult to practice sometimes it becomes inlet when you keep practicing and as soon as you look at a model you can see there is some structure anomaly. But to get to there you have to practice, so that is why we are playing with simple structures to look at this structure. We just start with first principles is that a first order system, is that a positive feedback is that a negative feedback flows in and flows out then only certain behaviors can be done if model is showing any other behavior. Then we can be sure that or we can at least have doubt it saying that no this does not look feel right.

And then once you get that feeling you have to act on it. Either you have to get more confidence in the model that what it showing is right or you have to update the model to conform that your beliefs. This is broadly what we can talk about as model validation.

(Refer Slide Time: 17:00)



Once you have verified an valid model one of the most common and popular thing that we ended up doing is called as sensitivity analysis. Sensitivity analysis computation of the effect of changes the input value or assumptions on the outputs Typically we start with very small changes or it typically change with one parameter value or one link or one assumption and see what is the impact of it within the entire dynamics.

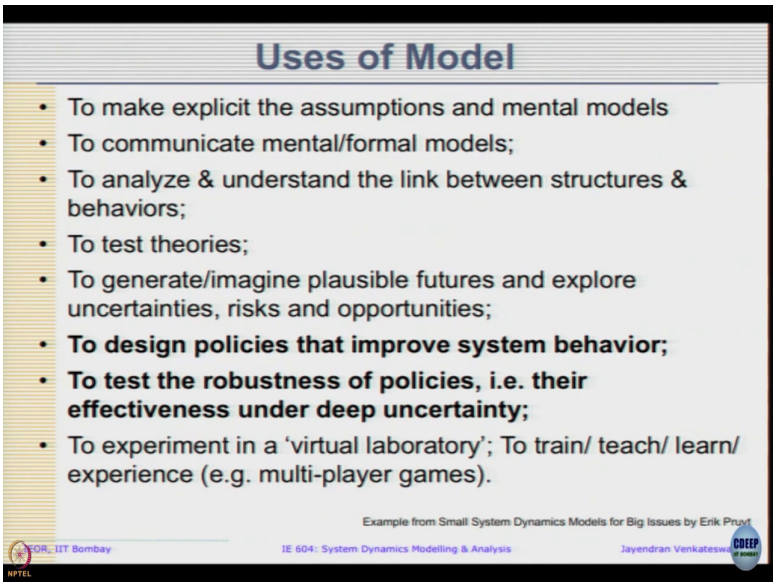
So, as to isolate the reasons or the affect of a single parameter changes. The change in assumptions whatever the assumption can be any parameter value or anything. If it results in only change in numerical value of the result that is called as numerical sensitivity if the pattern of behavior generate itself changes then it is called as behavior remote sensitivity.

Yesterday in the muskrat example when we changed the what is that proportionality variable we got exponential growth, or we got a exponential decay, or we got a stable parameters right.

So, that is a behavior mode sensitivity and that behavior mode is sensitive to that parameter value, so the entire behavior changes. But for certain parameters maybe it is exponential growth, but only the numerical value changed after few years may be, so then it is called numerical sensitivity.

If it reverses impact of proposed policy is called as policy sensitivity we look at policy analysis in a minute. So, this is broadly what sensitivity analysis is going to do and we will learn how to change one parameter time and say how we can document and compare the results to see whether what is more sensitive or less sensitive to changes that with our save.

(Refer Slide Time: 18:48)



This slide actually summarizes what is actual use of us doing all this modeling and model building exercises that we are doing throughout the semester and doing on is as follows. These are broad users why you are even doing all the model. It is to make explicit the

assumptions and mental models that we may have to communicate the mental or formal models.

To analyze and understand the link between structures and behaviors, to test theories to generate plausible futures and explore uncertainties risks and opportunities. To design policies that improves system behavior to test the robustness of the policies on their effectiveness under deep uncertainty. To experiment in a virtual lab to train, teach, learn, experience etcetera are all the broad purposes of model building.

The first few ones they seem quite abstract is when you self learn, but then you can for example, go to the last one right. So, using this models as a virtual lab you yourself can train games to understand what kind of dynamics is happening within the reality or a supply chain or a small factory or how infectious disease spread right. So, that will give an idea because we can gain real world experience in everything.

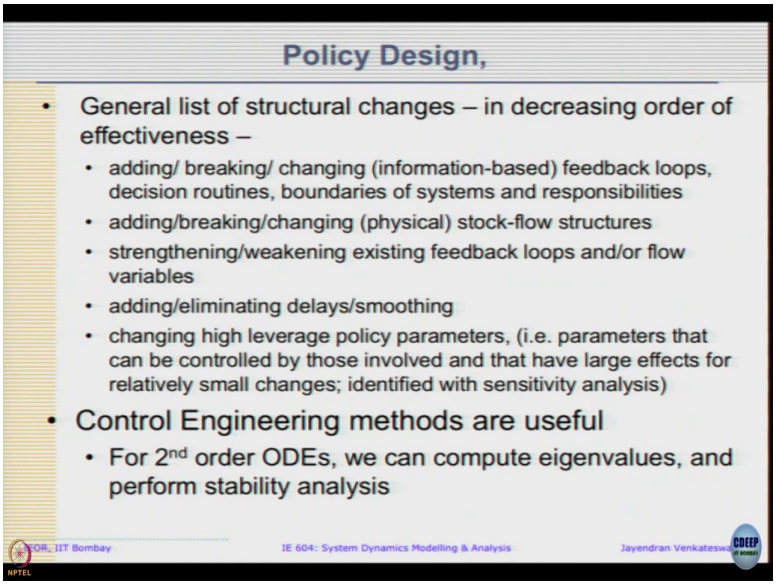
But it is important that you look at the model play with it and observe the reality in a better light. Like in campus I am not sure what happened last few years, but the decade ago conjunctivitis outbreak is common. Come monsoon season around that time there will be few case of conjunctivitis and suddenly it is going to spike up and then it is going to come down. And in fact so much that the conjunctivitis even the drug medicine there will be a small shortage. Because, so many people get infected which you can now understand much better because even a few people get it even the contact rate the amount the population density inside campus and hostels.

More people were eventually bound to get it before they actually seek treatment and then take precautions and without that it is not going to fall right. So, these are some simple models you can use to relate and observe what is happening in reality better, so that is what the last point just I trying to say. Whether two points highlighted in bold yes after building all these models what we actually want to do is come up with better policies to improve our actual system that is in place.

Tests the robustness of it see how effective it is when things are becoming so uncertain right. For example, consider infectious diseases if suppose our hospital wants to know how much to stock. And because all these unfortunate medicines come with the expiry date we cannot stock infinite amount when you do not have space its very expensive.

So, if you want to come up with some policies or to control those infectious diseases what will be the most effective policies? It is better to do kind of simulations and tests these policies. So, before we actually test the policy we need to design the policies then we test it or simulate it and then we analyze the policies.

(Refer Slide Time: 22:06)



So, policy design the first stage to be gain policy design we need to incorporate suppose we know the policy. Like for this infectious diseases the policy is if anybody gets infected they

will be quarantine or kept in a separate room, so they can avoid contact with people. Yes, is it a good measure or not this quarantine whatever we do is going to cost money?

It is a quarantine we need to ensure it is a whatever contact free room enough beds has to be provided you know facilities then people has to wear all masks this that. So, the contact is all minimized then when a patient is identified and put in quarantine if that would be better or should it give more vaccinations and awareness drives.

So, what are the different policies that can adopt to be very large? Which is the most effective one we can help uncovered? And when we incorporate any of these policy designs within a simulation model we end up changing the things in simulation model. Over years people have done various kind of policy analysis and in policy what we want to do we want to introduce some sort of structural change in the model right there is some decision making happening we want to introduce some structural change. So, decreasing order of effectiveness these are structural changes that typically helps in making high impact in your outcome.

The first one is adding, breaking, or changing information based feedback loops decision routines, and boundaries of systems and responsibilities. So, once you have model we know what is the physical flow and we want to know what is the information flows. Here it says that the by changing it by modifying that information feedback path we will have the maximum impact in your systems output.

The second one which can has the maximum impact could be adding, breaking, or changing the physical stock flow structures themselves. We really need to have so many levels of stock or is there way in which I can cut down one level within the system itself. Strengthening and weakening existing feedback loops under flow variables, adding eliminating delays or smoothing's changing, high leverage policy parameters that is parameters that can be controlled by those involved. And that have large effects for relatively small changes because you have identified them with sensitivity analysis.

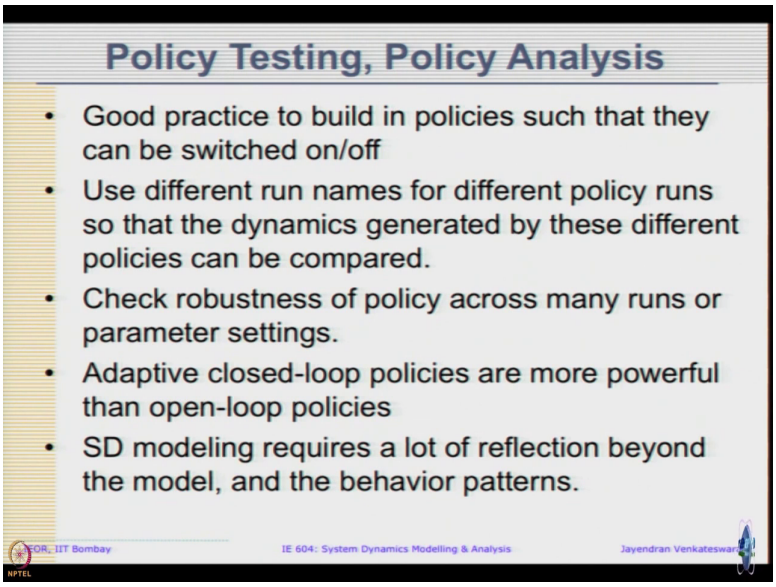
So, typically whatever policy analysis we do? These are typical things will end up affecting in our model. So, if you build the model we will be able to figure out that these are feedback

loops let us see where I can reduce this delay somewhere. Let us see where i can use this additional thing to make as a information and make a better decision within the model ok.

Or controlled theories is also very useful control engineering or control theory methods are also useful. See for example, the systems are all second order odes, so we can actually do computes eigen values and perform stability analysis. We can do what is it we can we can do z transform or discrete even system calculate the stability boundaries bounded input, bounded output stability analysis can be done. Which is bibo stability all those things will be useful even here because the underline system is differential equation or difference equation; however, you say it.

And whatever control systems ideas you have you can figure out what are the boundary conditions for the parameters which will results in stable behavior in system unstable behaviors oscillatory patterns can also be computed using the control theory ideas right. For example, if all the eigenvalue roots are positive then it you know the system is unstable if it is, so like that you can compute it. So, if it is negative then system is stable and so on. So, this is also form of policy design and testing where it looking at the extreme conditions.

(Refer Slide Time: 25:57)



So, once you have these policies we need to test it with a simulation model a good practices to build the policy in your models such that it can be switched on and off. So, it is a single model policy on, policy off you can have a variable. Please use different run names for different policy runs or dynamics is generated by different problems can be actually studied and compared. Checks robustness of the policy across many runs or parameter settings. Because, it is not just one comparison we need to compare across various settings similar to sensitive analysis for each policy.

Adaptive closed loop policies are more powerful than open loop ones. As soon as I start closing the loop system will more powerful than if you have open loop. So, whenever there is a policy session requested think of closed loop policies. And room where SD modelling requires a lot of reflection beyond the model and the behavior patterns.

Because, here it assumes that yeah yes real system is also quite structured and it is going to behave the way we expect it to behave. But reality may not work like that people may not even agree to these results and there will be strong resistance to changes. So, how to even incorporate the policy changes within reality you require more reflection beyond what is we can actually show.

So, it is a broad summary or outline for how to do model from yesterday's class model debugging then verification validation and then moving on to sensitivity analysis or policy analysis. It necessarily goes in that order because first we want to ensure that we have verified debugged model without any errors. And then its valid enough based on excepted theories or based on real problem that we have making. And once you are sure of that then we can do sensitivity analysis and policy analysis to uncover or understand the system better.