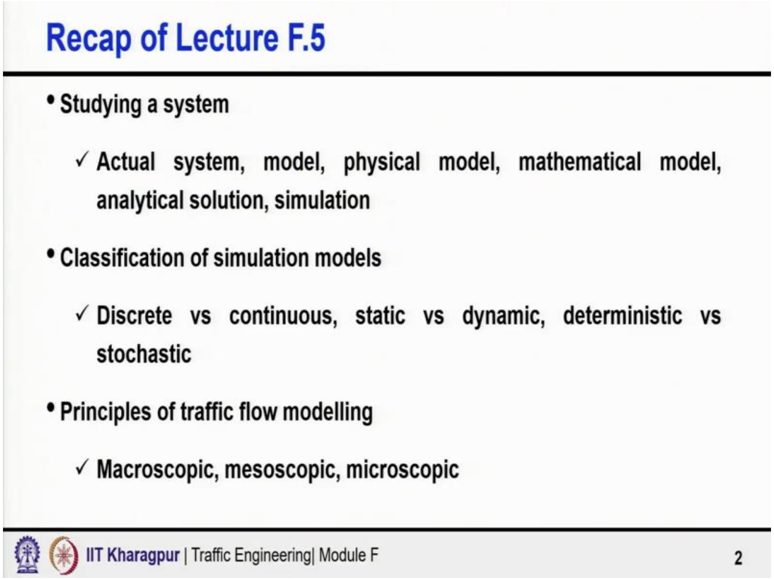
Traffic Engineering
Professor Bhargab Maitra
Department of Civil Engineering
Indian Institute of Technology, Kharagpur
Lecture 49
Traffic Simulation - II

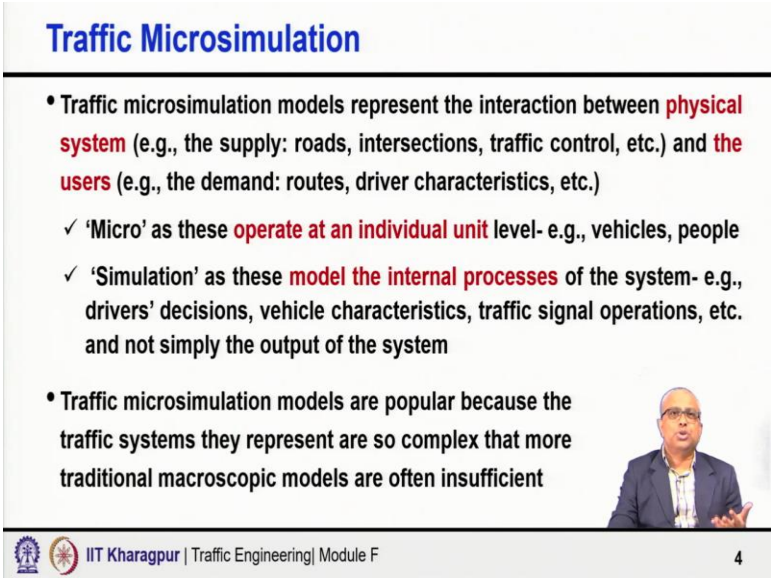
(Refer Slide Time: 0:16)



Welcome to Module F lecture 6. In this lecture, we shall continue our discussion about Traffic Simulation. In the previous lecture, we discussed about systems and the different ways the system can be analyzed, modeling or experimenting with the actual system, then physical model, mathematical model. Once you think of mathematical model then the analytical solution or simulation-based analysis, all those possibilities we discussed.

Then also the classification of simulation model, discrete versus continuous, static versus dynamic and deterministic versus stochastic analysis. Then the principles of traffic flow modeling, macroscopic analysis, mesoscopic analysis and then microscopic analysis and in lecture 5 itself I told that the next lecture would be on traffic simulation and specifically focusing on microsimulation or microscopic analysis.

(Refer Slide Time: 1:27)



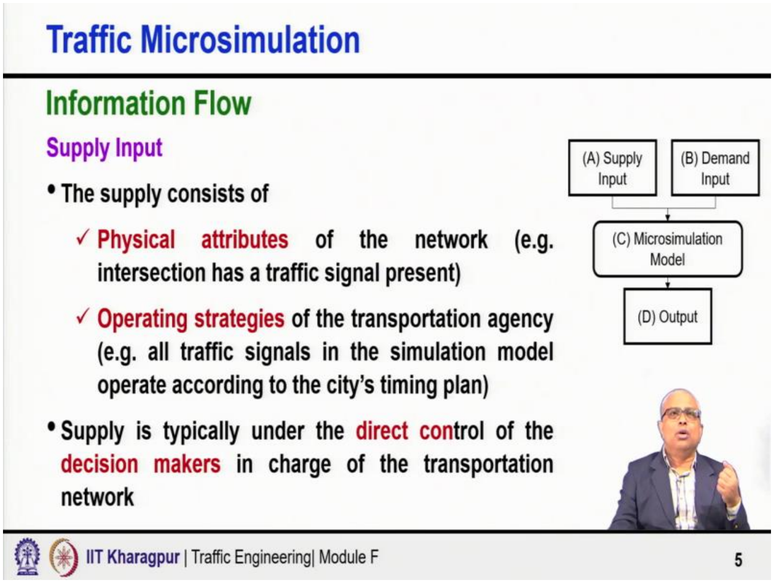
So, with that background, today, we shall discuss about the traffic microsimulation in fact, today's lecture and also the seventh lecture in this module, the next one would be on traffic microsimulation. To start with traffic microsimulation models represent the interaction between the physical system and the user or you can say it represents the interaction between the supply side and the demand side. So, the physical system of supply side may include roads, the physical infrastructure, the intersection, the traffic control aspect and the users or the demand side maybe present routes, driver's characteristics, etc.

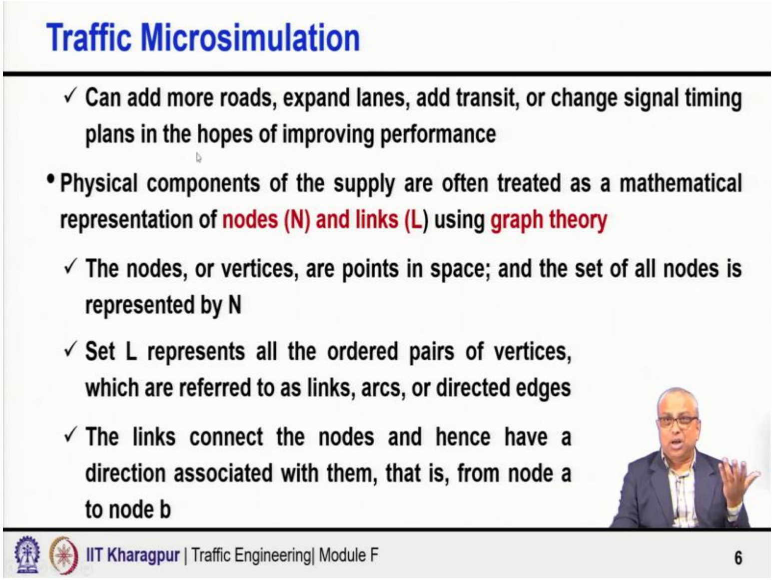
Now, while we are calling it traffic microsimulation. Of course, it deals with traffic, so, it is traffic microcirculation. But then why microsimulation, why micro? Micro because these operate at an individual unit level that means, vehicle, people that individual level individual entity level it operates.

So, the models are called micro models, micro level analysis and why simulation because this model the internal processes of the system. For example, the how divers are taking decisions, the drivers decisions are getting reflected, the vehicle characteristics are getting duly considered, the traffic signal operations are getting considered and not simply the output of a system.

The key aspect is basically the internal processes not simply the output. So, it is basically we are going for simulation and altogether we call it traffic microsimulation. Traffic microsimulation models are increasingly becoming popular these days. Most of the analysis we do we actually come out or take help of traffic microsimulation, microsimulation because the traffic system they represent are so complex that more traditional microscopic or macroscopic models are often insufficient. So, the macroscopic models cannot really capture well the complex behavior and it is desirable to use traffic microsimulation.

(Refer Slide Time: 4:19)





Now, how the information flow happen in a typical simulation model and of course, we shall focus in the context of traffic microsimulation. You have inputs to the simulation model. These inputs could be as we say that it represents the interaction between the supply side and the demand side. So, the inputs related to supply, inputs related to demand those are all going inside the traffic microsimulation models and we are getting our desired or target outputs. Coming to the supply inputs, the supply consists of physical attributes and operating strategies.

So, once you talk about supply inputs, we know what is supply that I have already defined. So, we say the physical system roads, intersection, traffic control related to that. So, the physical attributes within the supply actually it includes the network say for example, the intersection has a traffic signal present. So, the physical attributes of the transport network that is one side. The other side is the operating strategies of the transportation agency.

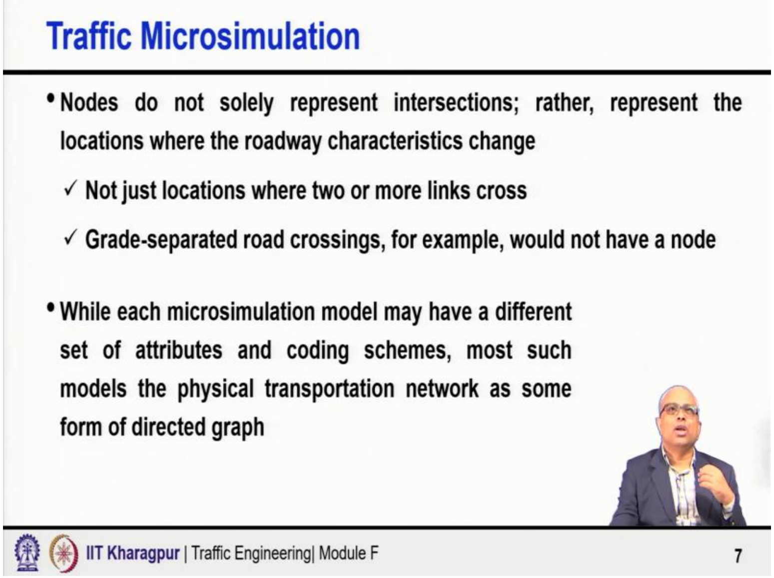
For example, while the physical attributes of the network include or tells us that an intersection has a traffic signal, the system or operating strategies will tell us information such as that may be all the signals in the simulation model operate according to the city's timing plan. That is the operating strategies. So, physical attributes and the operating strategies together become the supply input. Supply is typically under direct control of the decision makers in charge of the transportation network.

Because if we wish, we can change the physical attributes, you can actually install a signal in a given network, if you believe that, that is going to help us to improve the overall operation. So, that is what I said here that why we are seeing that supply is typically under the direct control of the decision makers, because we can add more roads, we can expand traffic lanes and more number of lanes at transit in the system, install signal or chain signal timing plan in the hope of improving the performance.

Now, physical components of supply are often treated as a mathematical representation of nodes and links. And here there is vast application or wide scale application of the mathematical graph theory. We often use it and we can check how the connections are there with that the network is connected with the what are the shortest path or whatever the kth shortest path and many such kind of things we consider and extensively the mathematical graph theory is used.

So, the node and vertices are points in space and the set of all nodes is represented by N. Similarly, the set of L represents all the ordered pair of vertices which are referred to as link, arc or directed edges. In traffic sense, we often call them as link. So, Link node diagram. The links connect the nodes and hence have a direction associated with them. For example, that node from a to b because direction is very important and the direction associated with the links direction associated with the links is important.

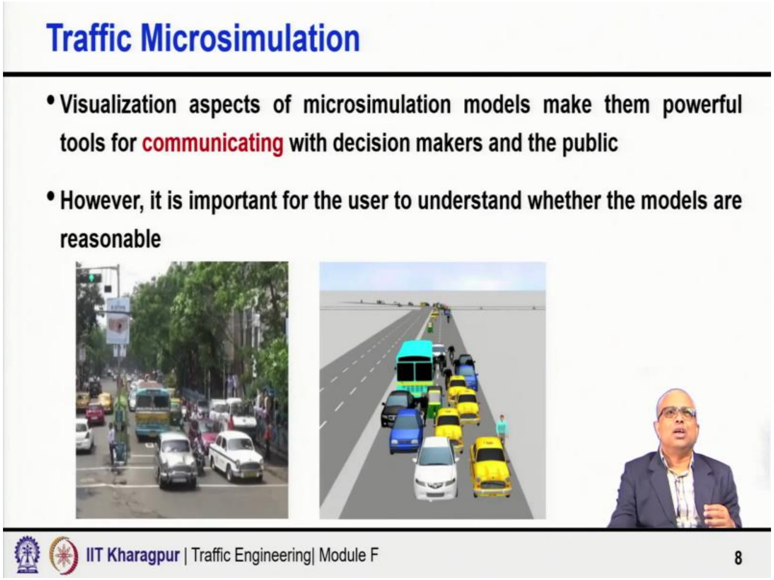
(Refer Slide Time: 8:51)



Nodes do not solely represent intersections; this is one important message or point to be noded by all of you. Not simply that two routes and meeting. Rather they present locations where the roadway characteristics change. For example, grade separated road crossing would not have a node because the traffic flow continuously. So, nodes are not just location where two or more road links rather the present locations where the roadway characteristics change.

While each microsimulation model may have a different set of attributes and coding scheme, because there are several micro simulators which are widely used all over the world by transport researchers, particularly those who are in the traffic engineering related works, but most such models, the physical transportation network is taken as some form of directed graph. That is some kind of thing which are mostly common across all microsimulation models.

(Refer Slide Time: 10:13)

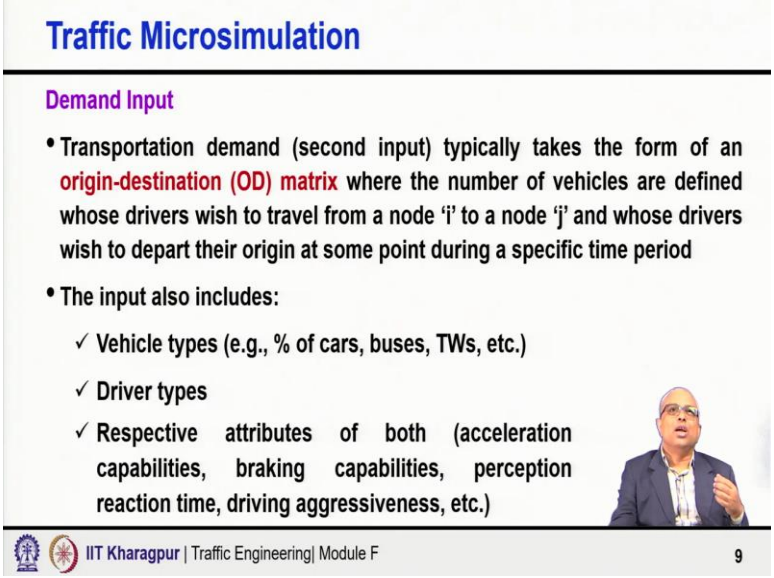


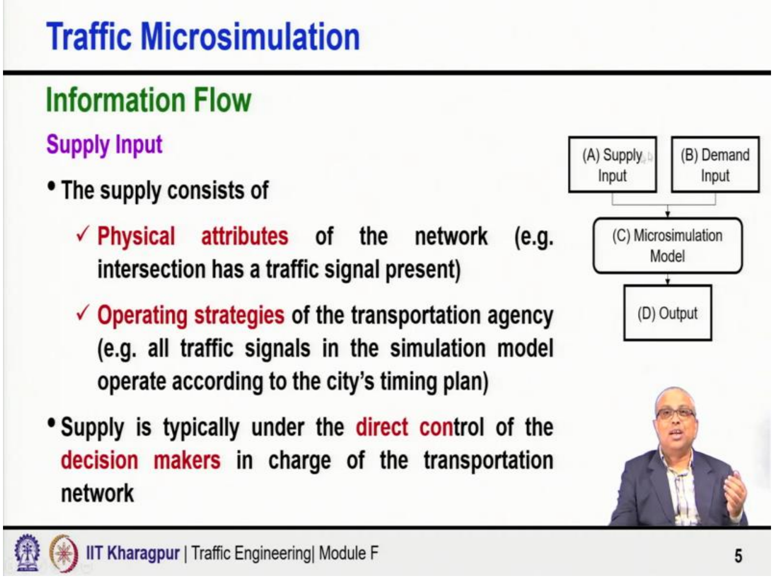
Now, one more thing is a very powerful tool simulation is a very powerful tool. And it has really the fantastic characteristics that the visualization aspect, it can tell you how reality how we really do is see the vehicles are moving, vehicles are taking turn or overtaking or doing the lane change, all such kinds of maneuvers can be represented and we can get, we can see the things visually. And that is one thing is fantastic. It gives us a feel. And also, we can see whether the behavior the way it happens on roads. With that we are able to replicate some of the similar behavior in the model or not.

So, visualization aspects of microsimulation model, make them powerful tools for communicating the decision makers, for communicating with the decision makers and the public. Decision makers, policymakers public, they get more convinced, if they are able to see, you say that if I have this change in the road geometry or the traffic control system, or have the bus priority, then the operation will happen like this. So, you show them that the microsimulation and with visualization and they get a feel. So, that is fantastic. That is really strict.

But remember that it is important for the user to understand whether the models are reasonable. That is really important. Visualization is strange, but it is important for the users to understand whether the models are reasonable. So, as I said the overtaking may happen, but whether the overtaking is happening properly or just simply you see that the vehicle is jumping. No, that is not really the correct behavior. So, it is important for us that to understand whether the models are reasonable.

(Refer Slide Time: 12:36)



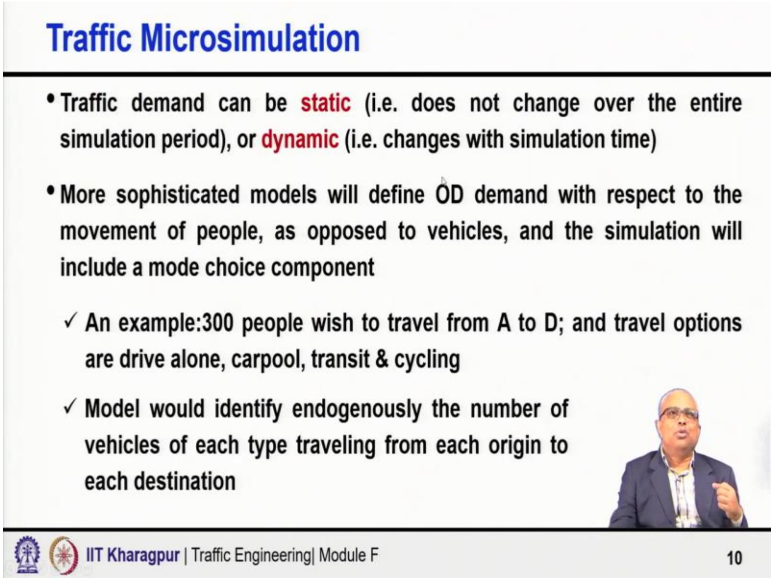


Coming to the next part demand input. So, what we said or discussed so far, about this supply input. Now going to demand input, demand inputs, most cases the transportation demand typically takes in the form of origin destination metrics. So, most cases, you may give actually, all the metrics. So, many people are or so many vehicles are traveling from point A to point B where the number of vehicles are defined, whose drivers wish to travel from a node i to a node j and whose drivers wish to depart their origin at some point during a specified time period, we are simulating maybe the peak hour traffic state.

So, that peak hour is from 8 to 10 or 8:30 to 10:30 or 9 to 11 depending on the characteristics of the city and accordingly you give the origin destination matrix. Now, the inputs generally include vehicle types, for example, you can say so many vehicles, but x1 percentage is car, x2 percentage is buses, x3 percentage two wheeler and so on, so forth. Also, the type of driver's, driver types and also the respective attributes of both drivers and vehicles.

For example, vehicle characteristics like acceleration capabilities, braking capabilities and driver characteristic for example perception reaction time, driving, aggressiveness. So, this could be the example. So, the inputs tells you the total volume, the vehicle type, driver type and the attributes of both drivers and vehicles. And often the demand is given in the form of a origin demand matrix, origin destination matrix, OD matrix.

(Refer Slide Time: 14:59)



Now, traffic demand can be static may not that means does not change over the entire simulation period, as I say that I might be simulating the traffic on the for the peak hour. So, I take the peak hour and within peak hour what is happening, I am trying to do their analysis. So, my demand does not change, I simply gave that peak hour demand is like this, during peak hour, this is my peak hour OD matrix, this is my composition. These are my characteristics of vehicles and drivers.

And I want to simulate the traffic to analyze different alternative scenarios, and to come out with the strategy that is most suitable for efficient management of traffic during the peak hour in a given network. So, here, the demand is static, because I am simply giving the peak hour demand. The demand could also be dynamic, when it is changing with simulation time, the entire period, dynamic traffic assignment we are doing. We are taking the time at different time slices. Maybe starting from 6 o'clock, we are going up to 12 o'clock to see that how the traffic starts at the very off peak state and then gradually increase and go to a peak and then eventually after some time come down.

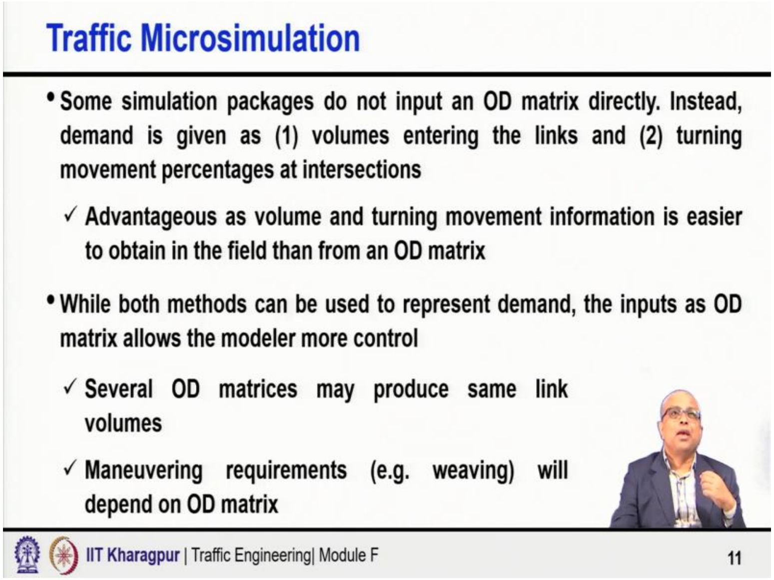
So, the entire range we are trying to simulate. And of course, here the demand at 6 to 7 or 8 to 9 or 10 to 11 they are not same. So, the demand is in that case dynamic it changes with simulation time. More sophisticated models will define OD demand with respect to movement of people. This is again another very important issue whether to take the demand as percent demand or vehicle demand. And the fundamental difference series, if we take the vehicle demand, we are only seeing given that vehicle demand how best we can organize the traffic.

But if you start with the person demand, then that gives another flexibility that you can see what is the effect of mode choice. If I improve public transport in the city, and then a public transport with the different services, how it is going to take the share of the total person demands. And remember that is the movement of people. That is our overall objective. Not really the movement of vehicles, vehicles come later. Ultimately, or eventually the person demands are to be converted to vehicle and demands. But then the mode choice analysis state driven in between.

So, that is what we said more sophisticated models will define OD demand with respect to movement of people, as opposed to vehicles. And the simulation will include a mode choice component. So, we will handle that part internally. An example maybe in that case, the inputs could be given like this 300 people wish to travel from A to D in that demand matrix. A to D there are many nodes in the network. Let us consider two nodes A and D just for example.

And let us say 300 people wish to travel from node A to D and the travel option given car driver alone using a personalized car, carpool or using transit like bus or even using cycle, bicycle. So, that gives you the flexibility and also an opportunity for you to see the mode choice experiment under different scenarios how the scenarios are getting impacting the mode choice and thereby the overall traffic state. Model would identify endogenously the number of vehicles of each type traveling from each origin to each destination that is being worked out internally. So, that is what I say endogenously, endogenously that will get worked out.

(Refer Slide Time: 19:59)



Some simulation packages do not input OD matrix directly as I mentioned earlier. So, far I mentioned that input is generally given in the form of OD matrix, but some simulation packages do not input the demand directly as OD matrix. In state, what they do, they give the demand as volumes entering the link how much volume is entering the link and the turning movement percentage at intersections that means, the if the left straight and right all three movements are allowed, then we will say 60 percent is going straight, 10 percent is taking right and 30 percent is taking left.

So, the it gives you the volume that is entering the link and determining movement percentages at the intersections. Both advantages and disadvantages are there. The advantages of giving such inputs that means the volume that is entering into the link and the turning movement instead of giving the OD matrix directly the advantages that volume and turning movement information are easier to obtain in the field then performing an OD matrix. So, if you want to really develop an OD matrix for a given network, even if it is small, it requires substantial effort, I would say.

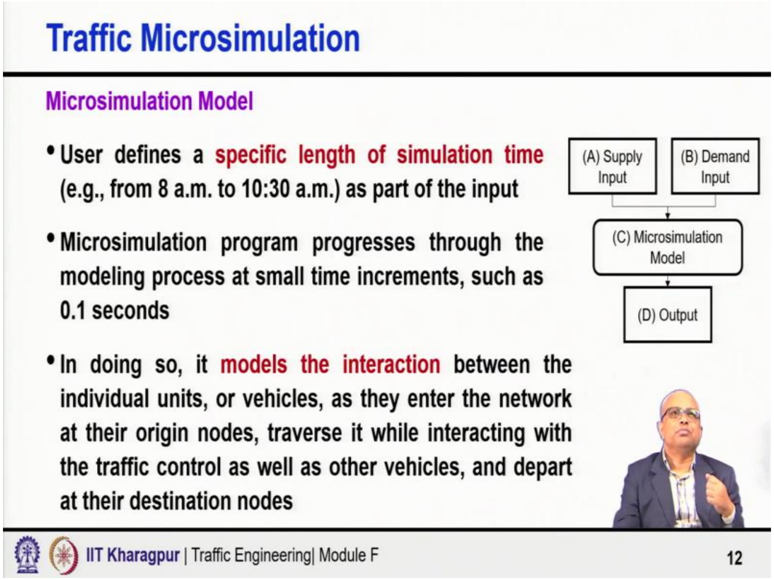
Rather it is very easy you just go and measure it, how much volume is there, just do the traffic counts and see that how the how much or how many vehicles are taking left turn, right turn and going through and then just approximately give the percentage. That is, field workwise getting the information that way it is very easy. You do not need to even captured the OD again and give as input.

But it has the OD has got the other advantages as well. So, both methods can be used to represent the demand but as they say the input as OD matrix allows modeler to have more control that has its own advantages. So, one way if I am giving the volume entering interlink and determining movement in percentage, then the field workwise getting the data is easy. But if we are using OD matrix, it allows the modeler more control why because you remember the same link volume can come from different alternative OD's. Multiple possible OD's, origin destination distribution may give the same volume on a given link.

And interestingly, the length characteristics, the outcome for example, how much time it will take, what will be the implications, conflict points, how many overtaking or say weaving maneuver is expected that will depend on the actual loading, not simply how many vehicles are going, but also the OD. So, the maneuvering requirements for example, weaving will depend on the actual OD matrix. And so, if you give the exact OD, whatever is the OD, if you give that then such kind of maneuverings will be captured in a more realistic manner inside the simulation model, and you are likely to get more realistic outputs.

So, it is not so easy to get the OD, but if you get the OD, modelers have more control. The maneuvering requirements and actually the effect of the level of weaving on the congestion or on the conflict or maybe the overall safety performers such kind of things may be evaluated. But as I said, if you give directly the volume and the percentage left and right turn and straight, then it is easy in terms of field data.

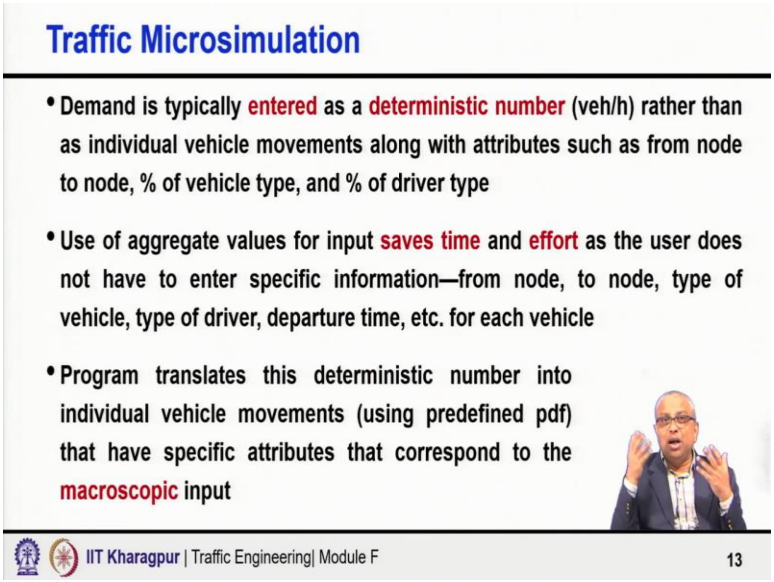
(Refer Slide Time: 24:49)



Then going to the see that is the microsimulation model. So, we discussed about supply input, we talked about also demand input. Now, we are going to microsimulation model. In microsimulation model user normally defines a specific length of simulation time, as I said that you can simulate only one-hour traffic during the peak or you can decide that how long it would be two-hour, one hour longer period of time as a part of input. So, you can decide as a modeler, what is the specific length of the simulation time, which period we are trying to simulate?

Microsimulation program then progresses through the modeling process at small time increments maybe every point once again. So, I have given my whole time period simulation time, but then microsimulation program will progress through the modeling process at small time intervals. And in doing so, it will model the interaction between the individual units or vehicles how one vehicle is interacting with other vehicle, how vehicles are interacting with the infrastructure and controlling controlled processes, all such interactions, it will try to capture as the interpreted network from their origin nodes. Also, as they traverse while interacting with the traffic control as well as with other vehicles and finally, depart at their distribution route. So, all such kind of interactions will be captured and the updates will be available at small increments.

(Refer Slide Time: 26:59)



Demand is typically entered as a deterministic number assuming that we are actually simulating it not that always it has to be like this. There are always a lot of variations which are possible, but assume that you are actually doing the peak hour traffic you are trying to simulate. As I say that a lot of problems traffic problems you find during the peak hour, some links are congested, some network some intersections are choke and so on so forth. And you want to analyze and see how network alternative network development or management plan can help you to ease the traffic problems.

So, here that you are doing only the peak hour. So, you get you give the volume that 5000 or 1000 vehicles is the demand for the peak hour. So, the demand is maybe in that case in such cases typically may be entered as a deterministic number rather than an individual vehicle movement along with attributes such as from node to node, percentage of vehicle type and percentage of driver type. So, what I am giving telling you that you give it as a total number 100 or 500 or 1000 as I say, give it as total number rather than giving every individual vehicle as an input.

So, two things are there, one is deterministic and also the static, not really the dynamic, that is one aspect and the did so, if you are analyzing given period, you just enter the number suppose as I said, one hour during the peak hour you are doing you tell that is what is my peak hour demand and give it as a deterministic number rather than telling every vehicle this will be the this is the vehicle type, this is the driver type, this is the acceleration-deceleration characteristics and so on.

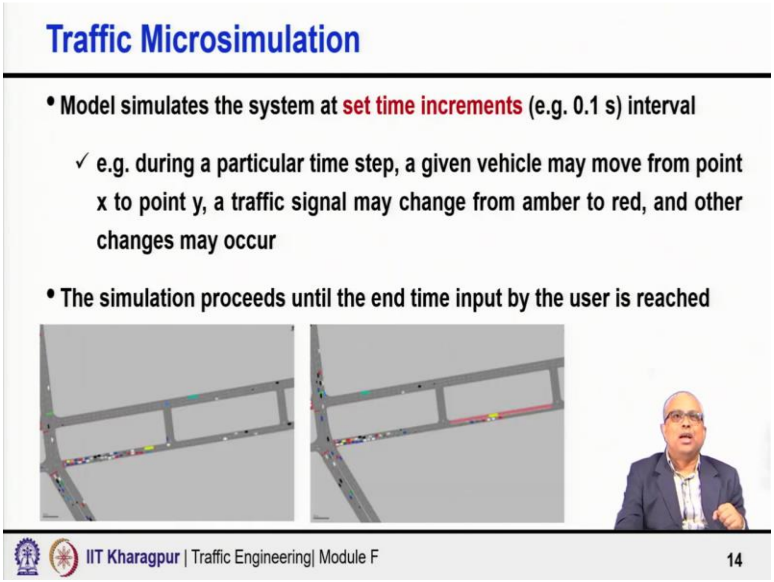
So, use of aggregate values for input it saves time when I give like a deterministic number the whole number have given it saves time and effort as the user does not have to enter specific information for each vehicle, where from it is entering, going to which node, what type of vehicle is, what type of driver will be assigned, what will be the exact departure time and such kind of minute details. You do not have to just keep the number. You can grossly tell what is the percentage of different vehicle type, what is the percentage of driver type.

Then the program inside the simulation model can translate this deterministic number into individual vehicle movement. Say for example, as I said that I want 100 vehicles during this period 100, 1000, 500 whatever it is let us say 100 vehicles. Now, if I say 100 vehicle I give it as a deterministic number but it will still generate inside this simulation model. It will actually generate the vehicle using predefined pdf. It may be vehicle or I am saying the headway, it could be the speed, it could be any other characteristic generally I am saying.

So, internally it will use if appropriate a predefined pdf and then generate. So, if you say 100 vehicles, it will actually generate also 105 vehicle, it will also generate 97 vehicles, both are possible. So, the deterministic number is actually inside the model inside the simulation model, the program translates this deterministic number into individual vehicle movement using the predefined pdf that have specific attributes that corresponds to macroscopic input. So, it will again get assigned that how many will be the 20 percent car, so, 20 percent will internally get assigned as car when the individual vehicles are generated, the speed it will get generated following a distribution.

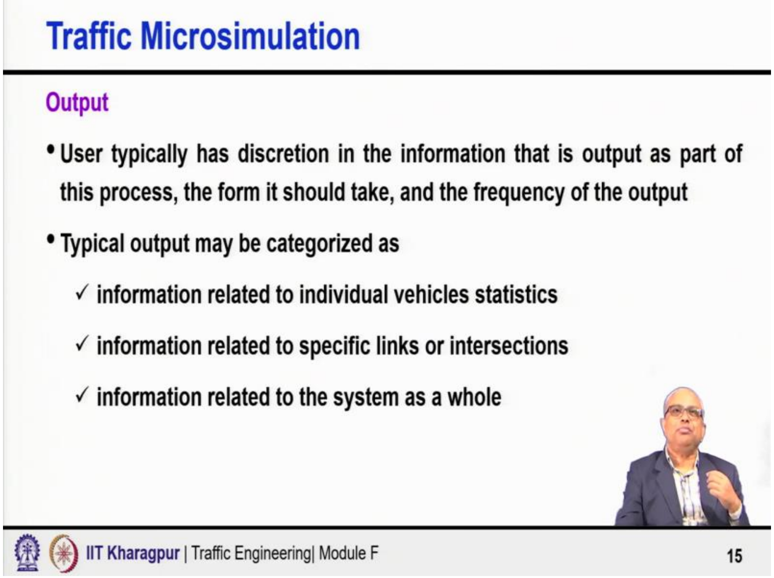
And similarly, every characteristic will get and even when they will start suppose you give this is the biopic period. So, they will start getting generating also following certain distribution. So, the input normally will give for easy because it is easy you give us deterministic number, but then it gets translated and all such internally this program translates this deterministic number into individual vehicle movement using predefined pdf and then attach characteristics to every vehicle.

(Refer Slide Time: 32:49)



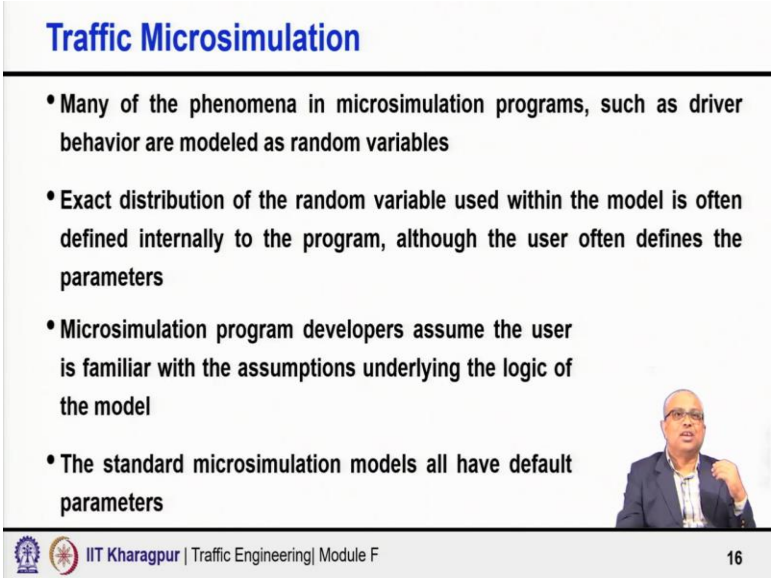
Model simulates the system at a set of time increments as I have said, say point once again interval A or so, and during a particular time step a given vehicle may move from point x to point y, signal may change from amber to red or other changes may occur and all these will be captured and accordingly updated the process will be updated. Simulation process proceeds like this until the end time input by the user as specified by the user is reached.

(Refer Slide Time: 33:34)



Now, coming to the output, User's typically has discretion in the information that is output as part of this process, the form it should take and the frequency of output. Now, the typical outputs may be organized as follows information related to individual vehicle statistics, information related to specific links or intersections and information related to system as a whole, aggregate level methods what we often say. So, all such kind of outputs may be of interest.

(Refer Slide Time: 34:15)



Many of the phenomena in microsimulation programs such as driver behavior are modeled as random variable. So, when we feel that that average value or expected value of whatever it is whatever you want to tell, actually many of the things are internally modeled as a random variable. I told you that that is why it is that you may give 100 vehicles but it will it may actually generate 97 vehicles, 98 vehicles or even slightly more than 100 vehicles. So, once you say that the speed follows these distributions and suppose it is normal distribution.

So, mean and variance or mean standard deviation whatever it is you give and then accordingly it will generate. So, the exact distribution of the random variable used within the model is often defined internally to the program. So, it is kind of inbuilt. But, remember that user often defines the parameter, what is mu, what is sigma that you can define. The parameters, which are actually controlling that distribution, which is rather which is describing the distribution those parameters you can give.

But often the model has the internal capability and all those distributions are almost inbuilt. Microsimulation program developers assume that user is familiar with the assumptions underlying the logic of the model. So, do not use it like a black box, many assumptions are inbuilt. So, you must know that when it is simulating and when you are getting the outputs, what are the assumptions that have gone inside and that are actually those what are the things that have influenced your outcome. And of course, in all cases, the standard microsimulation models all have default parameters.

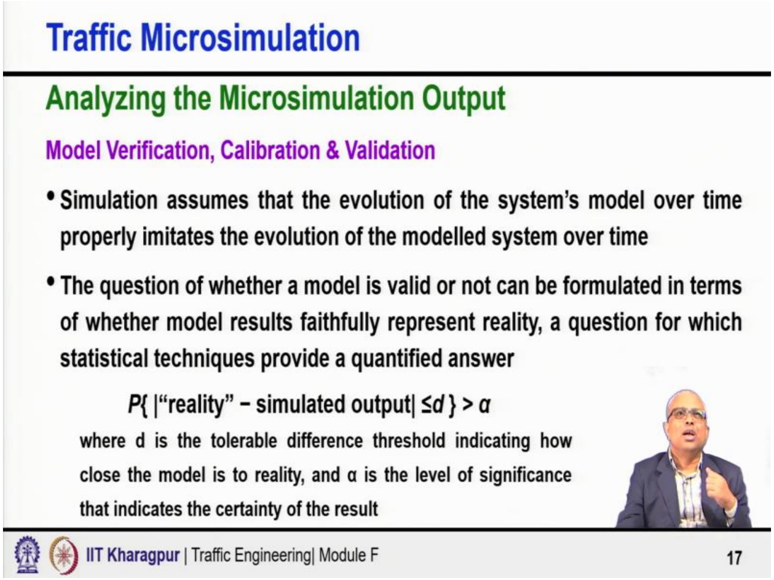
But here I would like to say something which is very important often the traffic environment based on which the microsimulator has been developed, maybe the default parameters are based on that environment. For example, maybe many of the microsimulator or the platforms have been developed considering the homogeneous traffic environment, more lane-based traffic environment, what we normally find in developed world or western world. Now, the default values might have been decided based on that environment.

So, if you are trying to utilize that same simulation model in Indian mixed traffic environment where practically except some on some elements, I should not say the lane concept does not exist that may be too harsh, but often the lane does not work because you have heterogeneity. So, much of heterogeneity, different dimensions, different characteristics, different culture I should say all together the lane concepts generally do not work. So, here maybe many of the default parameter values which are actually set as default considering a more disciplined homogeneous lane this traffic environment, if you try to use those values as default values, it may be very unrealistic in our traffic environment.

So, they are the question of calibration comes, you need to calibrate the model properly. But what is said here that some cases like the input many of the inputs, which are given common things like the speed distribution, the vehicle arrival or the counting headway or counting as appropriate, the distributions are inbuilt.

So, there it is not really complex if you want to assume that my speed follows normal distribution, I can always assume but that is what is said and it is said that you just simply give the parameters, the one to generate 100 vehicles, you give it, what is the mean headway, something like that or the total number you give. So, then internally the vehicle will be generated.

(Refer Slide Time: 39:35)



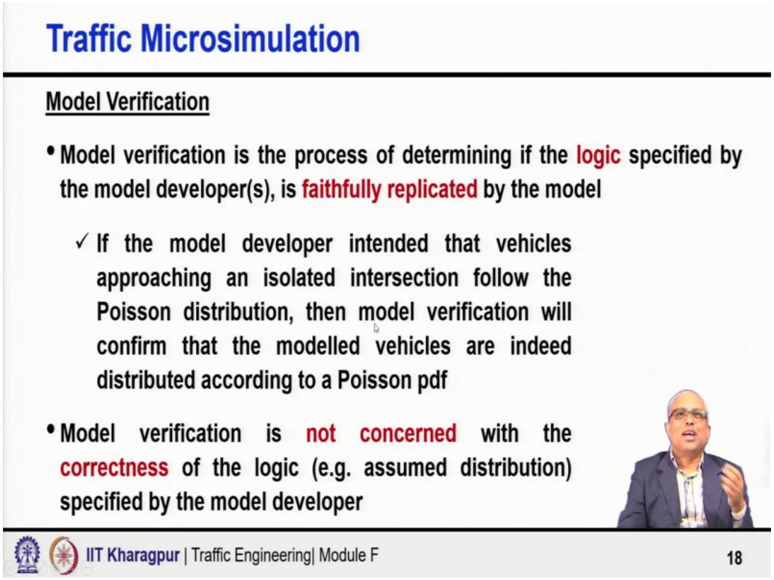
Now, analyzing the microsimulation output here the model verification, calibration, validation, all are important. Now, often you will find people are using such terminologies, but it is important for you to understand each of these terminologies very clearly. So, we will come to that. Simulation normally assumes that the evolution of the systems model over time properly imitates the evolution of the models system over time.

Now, the question of whether a model is valid or not can be formulated in terms of weather the model results faithfully represents the reality a question for which statistical techniques provide a quantified answer, if I want to really understand whether a model is valid or not.

$$P\{\ |\text{“reality”} - \text{simulated output}| \le d\ \} > \alpha$$

Basically, these statistical techniques can be used to provide a quantitative answer. So, what we are saying probability that the actual difference and or the magnitude of the difference between the reality and the simulated output is less than equal to d that means, in a way that my error is less than my acceptable error, d is the tolerance or tolerable difference threshold indicating how close the model is to reality. What is the probability that the error is within my acceptable limit? That probability should be greater than alpha. What is the alpha? Is the level of significance that indicates the certainty of the results that we can check? We will come back to this part later again and shall discuss. But now let us go one by one model verification, calibration and validation.

(Refer Slide Time: 41:49)



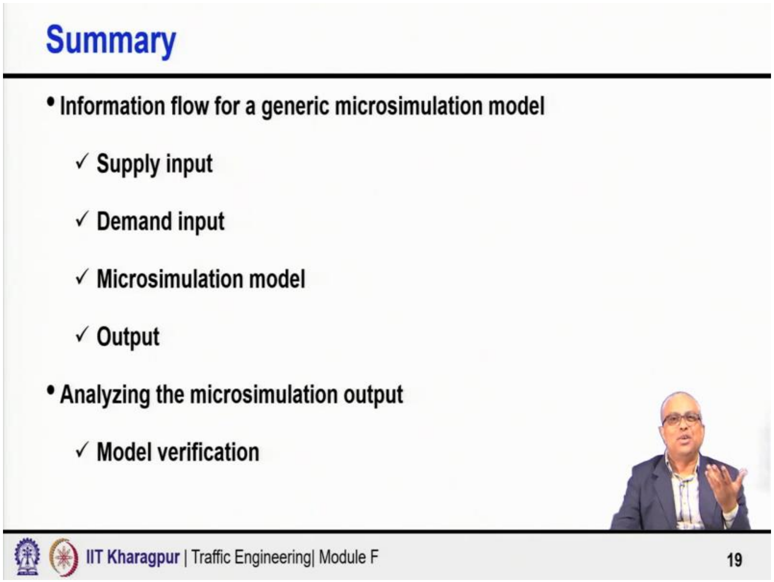
Coming to model verification, verification is the process of determining if the logic specified by the model developer is faithfully replicated in the model. For example, if the model developer intended that vehicles approaching an isolated intersection follow the poisson distribution, if that is what the assumption the modeler has made and modeler wants that to be implemented during the simulation process and accordingly picked up that okay take Poisson distribution as shown by the distribution. Then the model verification will basically confirm if the model vehicles are indeed distributed according to Poisson's pdf.

So, whatever modeler has achieved and wanted to happen whether that assumption is faithfully replicated in the model. I selected that follow Poisson distribution or follow normal distribution in some other context or any other distribution as a matter of fact. But actually, the values for the speed, number of vehicle arriving or whatever it is, with that the actual the what has been generated, what all have been generated in the simulation model, the generated numbers do they follow the same distribution, what the modeler wanted or assumed or wanted to be followed, that is what is a checking.

Determining if the logic specified by the model developer, whatever logic model developer specified, whether these are faithfully replicated by the model whatever numbers values model has generated, whether those are faithfully replicated, what the logic or the things whatever way the model is wanted it to happen. Model verification before I close very important, is not concerned with the correctness of the logic. Maybe you as a modeler assume normal distribution.

A during model verification, we are no way trying to check whether the assumption of the normal distribution is correct or not. That is not the thing what we are trying to do in model verification. All what we are trying to do you wanted it to be normally distributed with the numbers which have been generated, do they follow normal distribution? That is what is the model verification. So, model verification is not concerned with the correctness of the logic that means, whether the assumed distribution is appropriate for a given context that we are not going to judge. So, we shall continue our lecture in the next part.

(Refer Slide Time: 45:17)



We will continue about other aspects for example, model calibration, model validation and further carry forward the discussion. So, today what we discussed is general about the general information flow for a generic traffic microsimulation model and every discussion we made it contextual to traffic microsimulation, we said supply input about supply input, demand input, microsimulation model and output and then how to analyze the microsimulation output. And in that context mentioned to you only about one aspect that is model verification and the other things like model calibration, model validation and all related discussion. We shall continue in the next lecture. So, with this I close this lecture. Thank you so much.