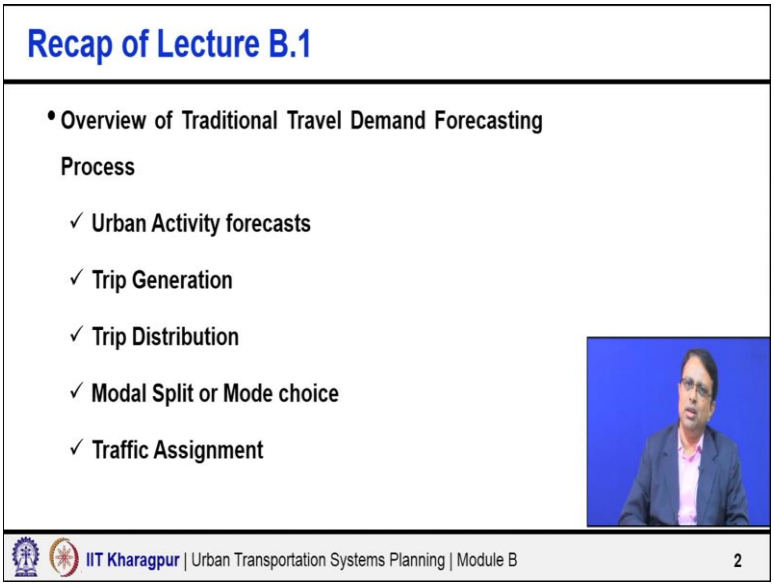
Urban Transportation Systems Planning
Prof. Bhargab Maitra
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

Lecture - 07
Traditional Travel Demand Forecasting Process, Specification, Calibration, Validation and Forecasting

Welcome to module B lecture 2.

(Refer Slide Time: 00:19)



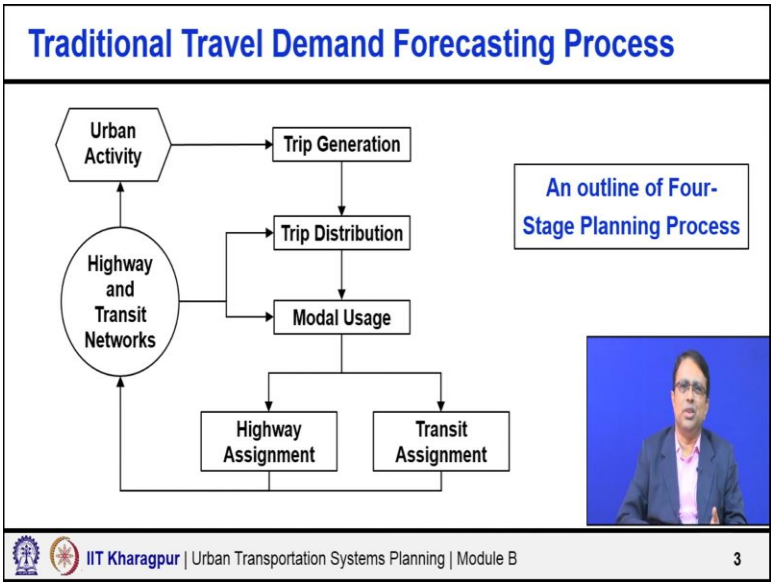
In lecture 1, we discussed about the overview of the urban travel demand forecasting process. We said that the urban activity forecasts provide input to this whole four stage planning process, then the four stages are trip generation, then followed by trip distribution, mode choice and traffic assignment. In trip generation, we convert urban activities into the number of trips. So the trip ends are obtained.

How much production or how much attractions are happening for each zone that we know trip distribution. We know the internal change that means if 100 trips are getting generated, where they are getting distributed or if 100 trips are coming to a zone, then where from those trips are coming. So we know each and every cell in the trip distribution matrix which are, row total and column total wise.

Satisfying the productions and attractions what we estimated in the trip generation process. Then third is the mode choice analysis. Here the person trips are getting converted into vehicle trips. Different modes of transport are available, private vehicles, buses, taxis may be sometimes even rail based mode. So how the vehicle trips are getting converted to different modes that we know and then finally the traffic assignment we know that point A to point B.

How many people are traveling by which mode and in traffic assignment, we know by which route that demand is getting satisfied. So how the routes are decided or how the routes are chosen. So with this we know that then how much travel demand will come at a given location at a given time.

(Refer Slide Time: 02:27)

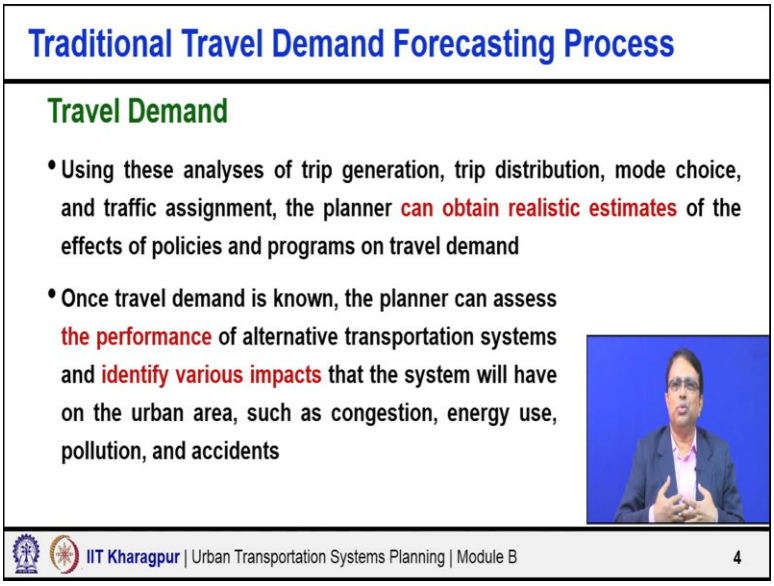


So that is what we discussed the urban activity forecast provides input to trip generation, then trip distribution, more choice and traffic assignment. Also we said that the transit follows a fixed route. So all bus routes are saved normally fixed route, but does not you know, change route all on a sudden. Whereas, the private vehicles depending on the road condition and overall network scenario.

The choice can be made by an individual. So the philosophy for assignment of transit and assignment of private vehicle the philosophies are different. So we are shown in different box and highway and transit network is very important, because the distribution road choices all are depending on the characteristics of the transportation system and finally how much loading are coming.

So, whatever assumption we make or we made during trip distribution and mode check analysis we need to check after assignment that how much really the load is coming in terms of vehicle volume and thereby how much really is the travel, time travel cost and accordingly we need to probably to update the distribution and mode choice if necessary. So that is the feedback arrow.

(Refer Slide Time: 03:51)



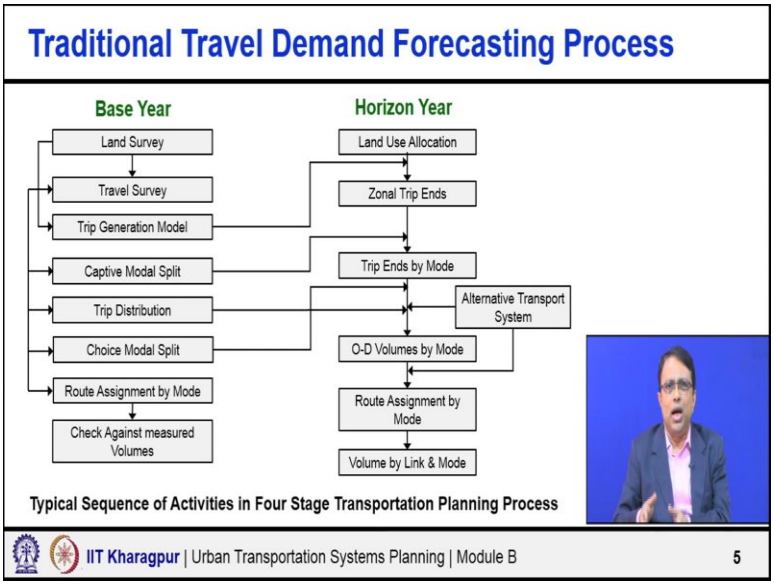
So altogether if you see the travel demand is very very important in the overall context of the four stages planning process and using the analysis of trip generation, trip distribution, mode choice and traffic assignment as I explained the planner can obtain a realistic estimate of the effects of policies and program on travel demand. For example, if there is a change in the transport network.

If there is a change in the fare policy of the public transport, or if you change the, you know, headway of service, how that is going to impact. That may impact the distribution, something may impact the mode choice, something may impact in the route choice decision. So any plant program policy you talk about how they are going to impact the travel demand that we can get through this four stage planning process.

And once we know the travel demand, we can know or we can assess the performance of the transportation system why? Because once we know the demand we know the capacity of each element of a transport network. Then we know what it means in terms of condition, what it means in terms of energy use and then thereby what will be the impact on you know air pollution and so on and so forth.

So we can also assess the impact. So under various programs various policies, what would mean finally in terms of congestion in terms of energy, in terms of pollution, even sometimes in terms of safety all these we can understand and thereby we can prepare well and we can say then, what is the right policy right action? So that, you know, the overall outputs or the externalities will remain within the acceptable limits or within the acceptable range.

(Refer Slide Time: 05:44)



Now, we show here now the actual application of the traditional four stage demand process. Something is called the base year and something is called as horizon years. What is base year? What is horizon year? Base year is the year we are actually doing the transportation studies or the year actually when we had developing all the models we are you know, building the models, that is called Base year.

That means in Base year our the independent and dependent variables both are known to us. Naturally, you know any model if you want to develop you want to develop a relation y equal to mx plus c just for example, you need to know y you need to know x, then only you can calibrate and find out what value of a and what value of a x plus b. So what value of a and what value of b will give you the best model and whether at all whatever model you have given.

Whether that is acceptable or not, all such kinds of decisions, you have to take right, one has to take. So we need to know both x and y when we are building the model. So Base year is that year where we know both input and output then only we can map. Say for example, you can say here we know that we do the land survey to know the existing how the activities are distributed, so the activities are known, we also through the travel survey.

So to know that actually how much travel is happening? Then the travel is also known. Activities are also known. So we develop now the relationship through the trip generation model and then try to capture that relation. So we do land survey, we do travel survey and with the knowledge of both land survey or the activity distribution and the travel survey we built this model what we call as trip generation model.

After developing the trip generation model, we can say that we do a captive model split. That means the mode choice, so far we said the choice of mode. But the choice of mode in general is available only to people who have got choice. So for example, generally let us say that public

transport of the bus system is the only available public transport. So the captive users captive means were kept to public transport that means they do not have private vehicles.

So those who do not have private vehicles, how they will travel, they will have to travel only by public transport. So there is no choice. So they are called captive riders. And those who have choice, who have got choice, those who have private vehicles they can either travel by private vehicle or they can travel even using public transport. So, we do the captive model split and then take the segment work empty to public transport.

So they will anyhow use public transport for them, the mode choice modeling need not be done. That is the whole purpose. Generally if there is only one common carrier transport mode or the public transport mode is available, say then we know that for them whether they will use public transport or they will use a car or private vehicle that is not actually valid, a reason to investigate well valid expect to investigate right.

Because they anyhow do not private vehicle, they are captive to public transport. So they have to use public transport. Then we do the trip distribution right, trip distribution also can be done earlier, then we go for the choice model split. Right the captive part we have taken to public transport already. Then the choice model split, so choice model split now we say given the public transport characteristics properties and also the private vehicle travel characteristics.

What share of these choice riders will actually go to public transport and what share will continue to use only private vehicles. So that mode choice comes there, those who have a choice and I repeat again who have choice, those who have vehicles, private vehicles. They have a choice to use either private vehicle or public transport for their travel. They are called model split, a choice riders and therefore we are saying choice model split.

That mean actual choice model happens here. Then we do the route assignment by mode because public transport is to be assigned separately, private vehicle trips are to be assigned separately the flexibility is much more for project you want to travel from point A to point B, you are the king you can decide which routes to follow among the available alternatives, but bus cannot decide in every trip, you know to follow a different route, that is not possible.

So the philosophy for assignment is different, that is why we say by mode and then end of the day we check against measured volumes, that means how good or bad we have developed our models in the base year as I say that x and y independent and dependent variable both are available both are known right. So we have the field measurement of the traffic right. So we can actually go back to the field, go to the field measure all volumes and then check how good or bad we have been able to model the whole system.

Say if we have modeled this slightly then whatever link volumes you are getting in the field or we are getting in the field should match with our model volume that means every road link from model also you will get some traffic volume, that if everything is fine as per my model I should say x1, number of cars, x2 number of buses, x3 number of you know, two wheeler and so on so forth on this particular link during the peak hour.

So you measure it also from the field and it should match give you a reasonable match. Do not try to match it number to number, every seat not necessarily. Please do not do that because you are doing a planning exercise, so the traffic changes whatever means you have measured today tomorrow you go to field and measure again in the same location, you will find some variations here and there.

The overall pattern will not change, say for example always peak hour you will get higher volume, but the exact number to number match will not happen. So do not try to match exactly by number. It is not necessary you might be distorting the overall travel demand models. So a reasonable match with 5%, may be even up to 10%, 5-10% accuracy as you know, error is also fine no issue.

So you can accept a model because you are doing it for the planning purpose. Understand that the dynamics of the data right, the variation of the data demand occurs if you go to any road section measure the traffic volume on consecutive 7 days, you will not get exactly the same number. So we are doing a planning exercise. There are many limitations for planning or the modeling also.

That also you should recognize we cannot be very precise in modeling approximately and a reasonable match that is what is the target. So we check against method volume and if we are not able to get the required match, then there are ways and means to know right where you have gone wrong you can think depending on what kind of mismatch you are getting. Suppose you are getting overall link volume fine.

But at the vehicle type wise match is not there, so good. So, you know, that means it may be it is a problem of then mode choice. But you are finding that you measured on all the roads which are top probably, you know, crossing a particular border or so you are getting overall match, but the route to route, you know, individual route wise match is not there. That may indicate that something you need to probably work with the traffic assignment model again.

So that means overall you have done. So similarly there are ways and means to find out where you have actually gone wrong. And accordingly, you can go back to the model and correct it. So all the model building development of the model calibration that means developing the equation

x and y both have to be known, if you want to develop a model and trip generation, trip distribution, mode choice, traffic assignment everything, you know.

You check against the measured volume in the field to get convinced that yes, whatever I have done. I am able to, get a reasonable match with the field observed traffic. So my models are fine. Reasonable match, again I am saying Reasonable match. So you do this whole thing. Then why we have developed all this model? Have you developed it for the Base year? Base year, you can go to the field and measure it.

So any modeling we have done or every model we have developed, we have developed because we want to apply this model for the future. So that future is called the Horizon here. Normally, you can select suppose ten years you want to do the forecast. So maybe you can say after five years as Horizon year one and after ten years as Horizon year two. That means if 2020 I have is my Base year.

Where I have collected all my investigation field data collection, model building exercise. I have done then maybe 2025 may be my Horizon year one and 2030 may be my Horizon year two. So in Horizon year, what do you want to do? We have to apply this model to get the forecast volume. That is why we have developed the model. We want to apply it for the future. So the application comes in the future.

How we are applying? Let us follow that flow chart. We are saying first the land, use a location carefully known. We are not saying the survey. For the Base year we said survey that means we are actually measuring it. Maybe secondary data also you can use but you can also measure many of the things, if you wish or if it is possible within the scope of the project or the work. But for Horizon year it is allocation.

Because when you are doing it in 2020 or 2021 and trying to see what may happen in 2025 or 2026 after 5 years or so then future land use you do not know you cannot go and measure the future landings. But there are planning authorities who have their plans and programs, they know how the overall area should develop, every development authority will have a vision how they want the area to get develop.

So you can actually know from those and you know the land use allocation. What is likely to be the allocated land use in the future? Once you know the allocated land use then you have already developed a trip generation model look at this box here. You have already developed a trip generation model. So then apply that model. You know that landings allocation, so your x is known and your y equal to this function of x that model or the equation is also known.

So you get the one what is the y, y is the zonal trip ends, because you have applied to trip generation model. So the zonal trip ends are known then you are also you can do you can develop treatments by mode. How I am saying, you know may be the, from the captive model split that you know, 50% population just for example, only 50% population or say 60% population do not have a choice for private vehicle.

So they are straight away you say them they will use anyhow in the public transport. Then the remaining 40% those who are choice riders, they may use public transport or they may use even private vehicles. So you would then, you know, you apply your trip distribution and also you apply your mode choice model and you give the alternative transportation system. This alternative transportation system in do nothing scenario may be the transportation system what is existing today.

So, 2020 or 2021 transportation system may be an input. Also if you think, no let me see if that corridor gets developed further what the government or the development authority is planning then you include that in the future network alternative transportation system. So whatever all transportation system you give with the transportation system for the developed distribution and mode choice model you will get the O-D volumes.

Origin destination volumes or the matrix by mode. So any it may be for do nothing scenarios, it may be just today's network it may be a future network development whatever you want. So you get the O-D volumes by mode, then you do the, you have already developed the assignment model, so you can get the route assignment by mode and you also get then future what will be the volume by link or by mode.

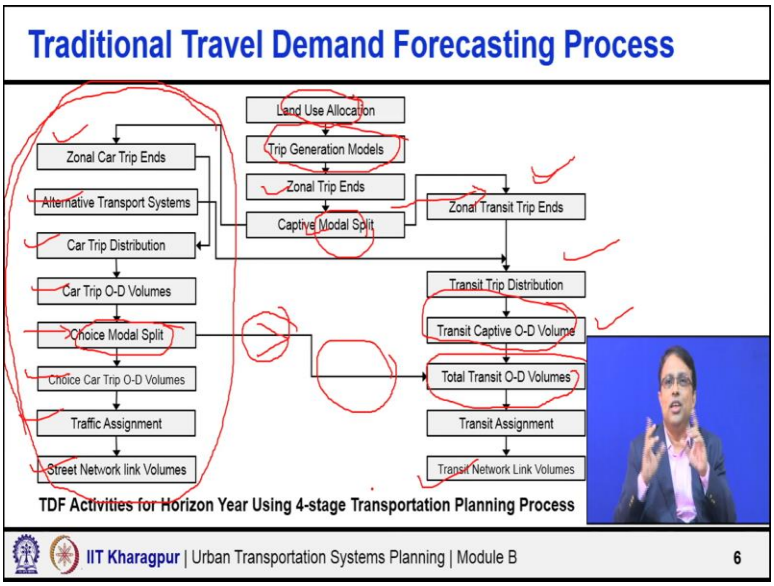
You take a link it will tell you how many vehicles will be there, how many will be cars, how many people you are going to use bus system, how many turning where may intersections, how many traffic for, how much traffic is going to use the intersection, what are the different movements, how many will go straight, how many will take left turn, how we will take right turn all such kind of things one can get right.

So for a given future scenario transport network scenario and future allocated land use, you applied all the models what you have developed in the Base year for the Horizon year to understand future volumes on every link, every intersection, public transport system you again, if you want to change the public transport system that may be also done as a part of this alternative transportation system and then you know.

Suppose you have changed something in the public transport network, then how does that influence the public transportation? And because of that how much reduction is expected in the overall private vehicle volume and thereby the condition, emission and everything. All the things

get linked. So this is broadly, how the way the Base year model is developed and how the way and the way it is applied for the Horizon year.

(Refer Slide Time: 21:58)



Now in this flowchart, I show little bit in more details about the application of the model in horizon here. So whatever I have shown earlier, the left part was the Base year and the right part was the Horizon year. Now the Horizon year application, I am showing in little bit more details. For better clarity, what we are doing? We are saying here we are starting with land use allocation. So let me take a pin that would be good.

So we starting with this land use allocation that is my input, then I am applying my trip generation models, which I have developed in the Base year. So I am applying that with input as the land use allocation. I am getting the journal trip end that is what is here. Then you do the captive model split. So once you have done the captive model split you got zonal car trip ends. Out of this car trips some of them still may go to you know, public transport.

But the initial thing is how many have car and how many do not have car. So those who do not have car, they are going to the right side journal transit tripping. Then let us come back to the left vertical. You have the zonal car trip end then you see the alternative transportation system, the future network, future transport scenario with widening of road, without widening, with a new link, without any new link.

Anything you want to give, you give it for the given future, for the future, you know, transport scenario. There you get that car trip distribution if that is the network then how the car trip is likely to get distributed. So trip generation has been done here, you have done here the trip distribution. Then once you have done the trip distribution. You know, the car trip O-D volume, O-D wise.

I to J, every I to every J how much travel will happen or how many people will travel that is known now. Then on that we are applying now the choice model split that is what it is. So you are now applying the choice model split. Because as I said all these are actually I started with zonal car trip data. So depending on the transport system how good or bad? Some of them or many of them probably will use or may use the public transport.

So you apply the zonal choice model. Those who will after models split indicate that they will use the, you know, public transport you take them to the right side. So here I will leave it now for the moment and the remaining people will continue to use, you know, car trip. So you have the choice car, trip O-D volume, then you apply traffic assignment and you get straight network link volume.

So this whole left side thing is now for the private vehicle end of the day, you know that how car trips are going to be distributed and those who are going to use the car also continue to use the car finally how they will do assignment is to order distribution is likely to happen and then how finally you how much you are expected to get in terms of street network, you know link volumes for cars.

Coming to the right side you already have this zonal transit trip ends from the captive modal split. Then you apply here also you give the alternative transportation system network, that may be also include some change in the public transport network and so then you get transit trip distribution. After that you get Transit Captive O-D volumes, why I am saying Captive O-D volume.

Because so far I am working only with the captive trips those who do not have option for car. Because I got the captive model split here and then taken only the transit part. Now this left side get linked with the right side. I also know after this choice model split for the choice riders or tripmakers, some of them will use the transit or public transport, so that is getting added here. So now with Transit Captive O-D volumes.

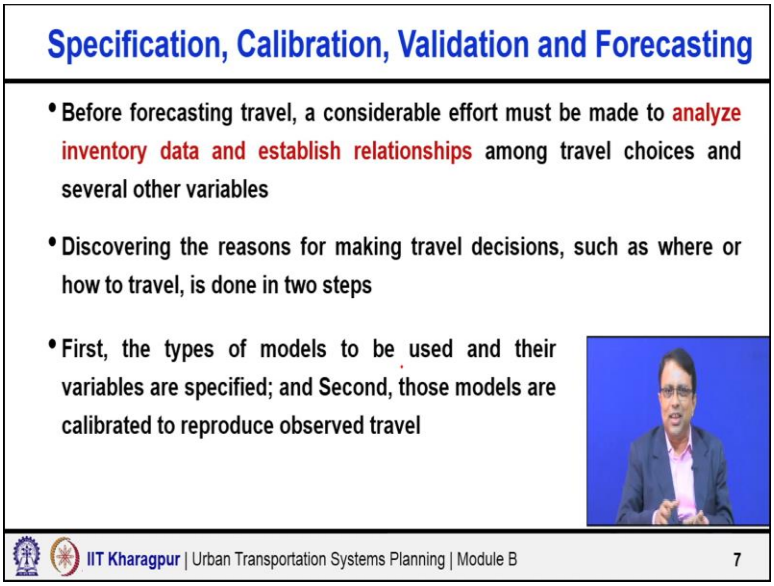
Those anyhow who do not have any option plus segment of people who are coming from the left side that means they have cars but they want to also use public transport. So now you get the total Transit O-D volumes. So Total Transit O-D volumes include for captive riders and also a segment of the choice rider will actually likely to use public transport. Then you do the transit assignment and get the transit network link volumes.

So end of the day the right vertical gives you how Transit Trips are expected to happen, how many people will use, what will be the you know in each link or each route or you know, each line what will come you know, all this and the left side gives you the other than transit the private vehicle distribution and you know, how much will come from that in every link or

intersection? So all together if you add them, you know that total picture in a transportation network.

That is likely to happen for the given transport network and given allocation of the land in the horizontal area.

(Refer Slide Time: 27:59)

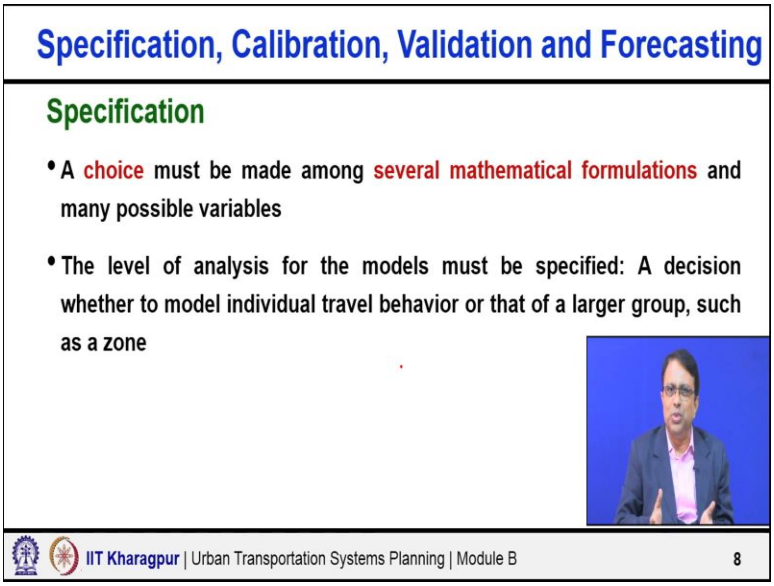


Now before I close quickly we will discuss four topics one is called specification, calibration, validation and forecasting. What is specifications? Specification is you are just you specify the model. You just simply specify the model, y equal to mx plus c, yes, it is a model specification. I do not know the value of m. I do not know the value of c, but I have specified a linear model. So before forecasting travel a considerable effort must be made to analyze inventory data and established relationship among the travel choice and several other variables.

So discovering the reasons for making travel such as where and how to travel is done in two steps, first the type of model to be used and their variables are specified. Then second those models are calibrated to reproduce observed travel, that means I have y, I have x, I have specified a model which is y equal to mx plus c. Then I will see for the given data if I try to fit this model, that means you calibrate, first you have specified that y equal to mx plus c.

Then you calibrate to get the value of m and c, which base describe this relationship for the given data that is the calibration.

(Refer Slide Time: 29:27)

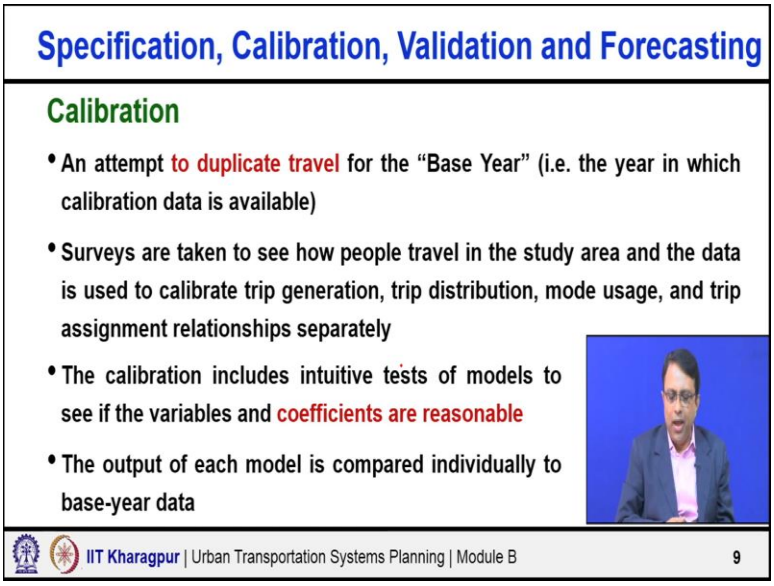


So first is the specification. So in this specification a choice must be made among several mathematical formulations, and as I gave an example of y equal to mx plus c, it is one of the many possibilities. You can have different, you can think of log you can think of power, in the so many different mathematical function, you can think. So the level of analysis for the models must be specified.

Say for example, you want to do like a zonal level model or you are doing a disaggregate level you are at individual level. You are trying to say the behavior that how an individual is likely to choose the mode right or choose the route. So what level you are doing, you are doing it aggregate level, gross level or you are doing at a disaggregate level or individual level or type of travel behavior model what typically you are trying to do.

So level of analysis and the type of mathematical relationship that you are assuming. So that is what is the specification?

(Refer Slide Time: 30:28)



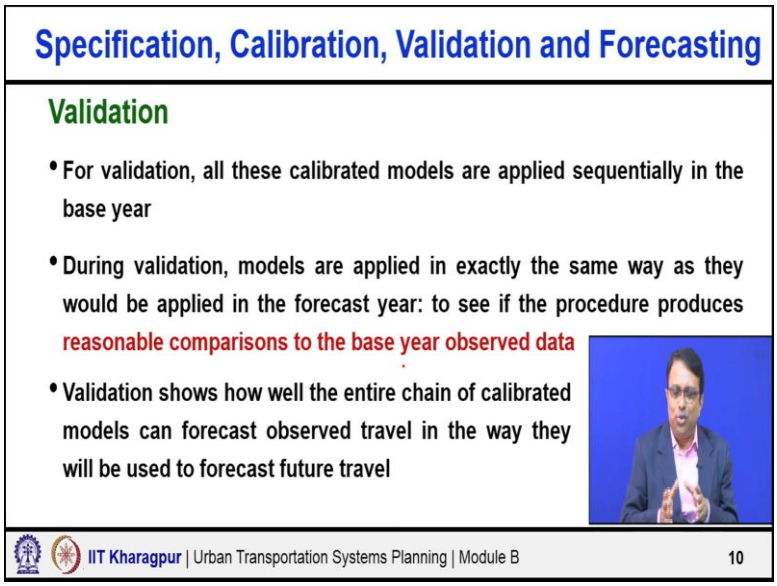
Once you have specified, then you are calibrating it; that means in the Base year, you know the y, you know the x for the, you know specified model how best the coefficients may be estimated. So that the relationship model specification in a best possible way described the data. Still the model that model may be acceptable may not be acceptable due to all depend on the goodness of its statistics.

How good or bad the data model is able to, you know, explain the variation in the data. But first thing it is to be specified, then it is to be calibrated, then only after calibration, you know that whether the specific model or the specified model is able to or is a good choice for that given data or given context or not, right. So an attempt is to duplicate here in place of calibration for the Base year data, now services are taken to see how people travel you have all the field data.

And, then the calibration includes an intuitive test of models and to see if the variables and the coefficient that are reasonable. So you may find a coefficient estimate m and c. But you may find c is a constant is a very large constant is a huge amount or large portion is actually unexplained. Then you know that my model may not be a good model or you get somewhere multiple variables are there where you are logically should get a positive sign you get a negative sign.

So you will not accept that model because the model does not you know is not logical right. So you have to check all the reasonability, but it is all part of the calibration. The output of each model is compared individually to the Base year data.

(Refer Slide Time: 32:19)

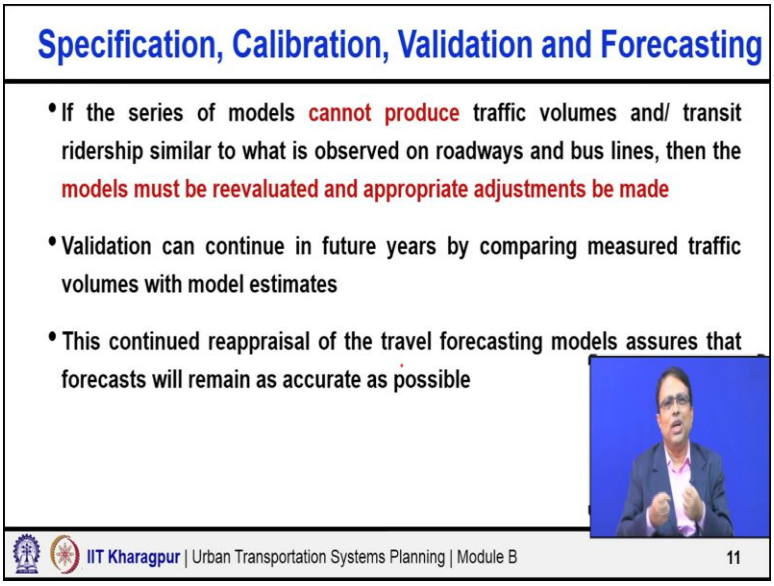


Then third you go the validity, you have calibrated it, you have now, you have to do the validation. How we do the validation in this context there are you know, many ways to do model validation. In transport contextual in the four stage planning, context this validation are done in a particular manner, which is acceptable. So what we do all these calibrated models are applied sequentially in the Base year.

That is the first thing sequentially we are applying to generation the trip distribution, then mode choice, then traffic assignment once all these models are applied, you now get the demand how much will come in each route, each intersection. If you take a screen line that means if you just simply take a hypothetical line from one edge to another edge it will cut maybe a few roads then you measure the traffic volumes there at all those location.

And you say in my model how many, how much demand I am getting you know in each of those locations which have is in the field if I measure how many or, how much demand I am getting in the field. So after applying this model in sequence the final, you know, distribution of demand whatever you have got in the network, you compare it with the field thing, field observers, and then if there is a reasonable match, you say I have validated it. But if it says shows that, it is not correct then you go back, so that is what is the validation?

(Refer Slide Time: 33:58)

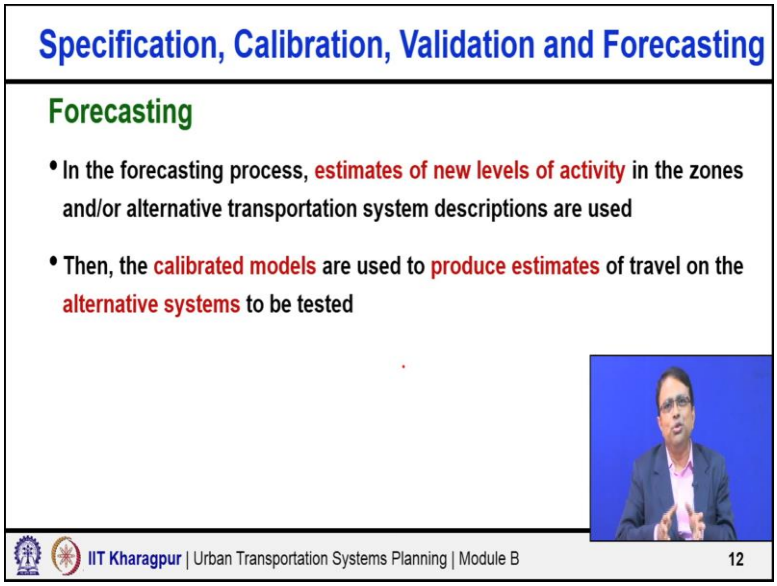


So if the series of models cannot produce the traffic volume of the transit ridership similar to what is observed on roadways everything is happening in the Base year, then you must reevaluate and appropriate adjustment you should do. So validation even can continue in the Horizon year because you know, we develop a model in the Base year then we apply it in Horizon year.

But for Horizon year when we apply we still have a traffic data, we may collect in a limited data, so you want to still validate the model. So validation may become even a continuous task. In many cities suppose a particular city you are working for a long, over a long time right may be most academic institutes also do like this. So maybe one IT is working or an IT is working on a particular city or a particular nearby town.

There, they have the data. So every time you get a new project you do not do the city wide data collection, but maybe little bit some data collection you do, so you simply validate your model and if you require some fine tune or some finer adjustment due to.

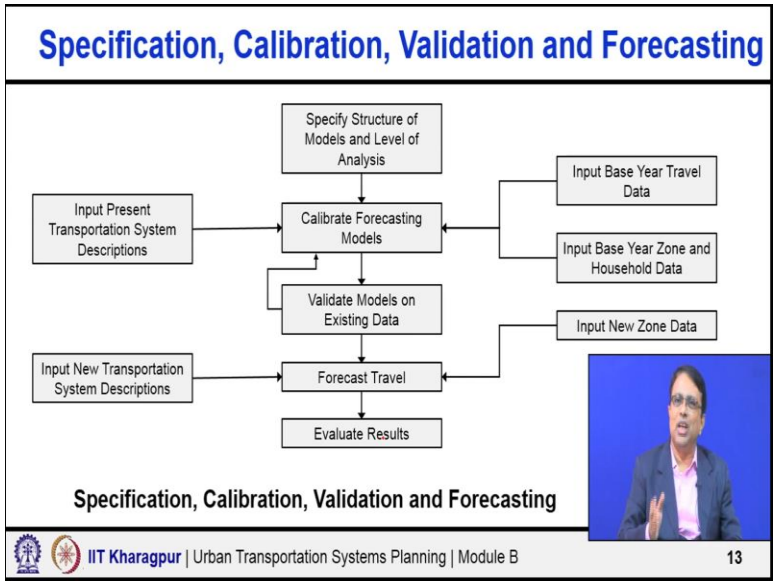
(Refer Slide Time: 35:08)



So that is what once you have the validated, then the calibrated and validated models. Now you are using those models for the forecasting in the horizon year. So first specification, then calibration in the Base year, then in four stage transportation planning context or hosted modeling constant series of models what you have developed you apply in sequence and finally whatever you are getting from the model in terms of link flows and all you know outputs.

You match it with the observed outputs and get convinced whether they are fine or not, if not then you do some further readjustment, go back to calibration, find to in a model which is actually giving you the error once you are satisfied in the validation is done. Then you use that in the, for the forecasting purpose.

(Refer Slide Time: 35:59)



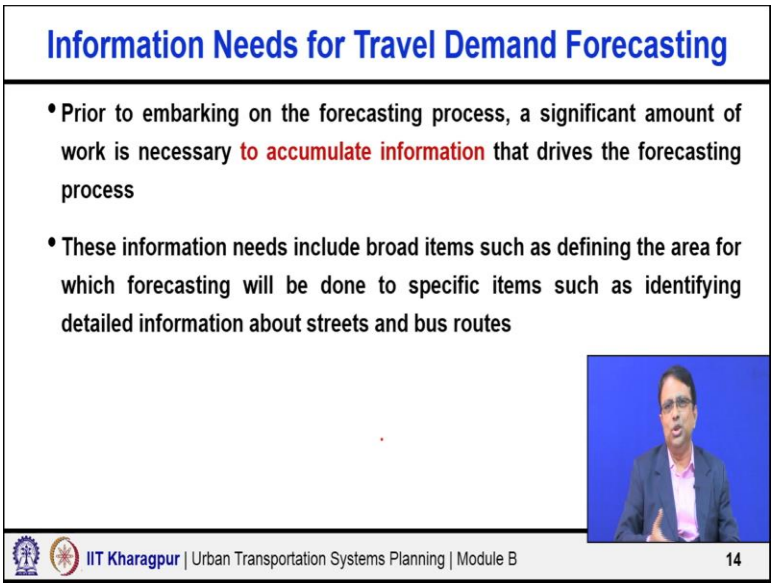
Finally I am showing this, you know, a simplified representation of specification, calibration, validation and forecasting here, you can see you are specifying the model structure, then you are

calibrating the forecasting models, when you are calibrating you need input of the present transportation systems. You need input for the Base year travel data, you need input for the Base year zone and household data.

All these are required for developing all these models Base year model. Then you are doing, for the going for the validation based on the field checking right model and field values after applying all those models and sequence then if you are not satisfied you may go back to again calibration and this process may repeat in a loop unless you get your satisfied with the validation. Then once you are satisfied with the validation you go for the forecast.

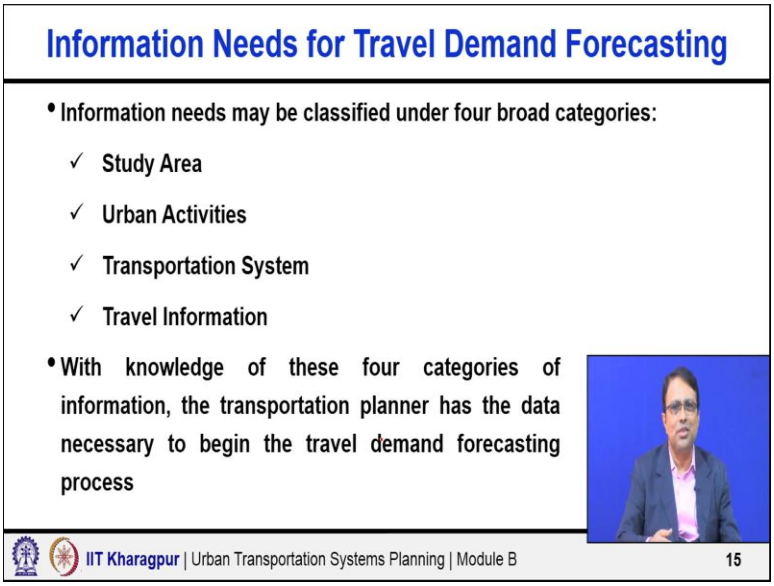
When you go for the forecast you use all the series of models, but you use future transportation system description and also the new zone data because that is the allocation starting from the beginning. And then you apply this model to forecast travel and then you evaluate the result. So that tells you clearly how different the specification, calibration, validation and forecasting.

(Refer Slide Time: 37:17)



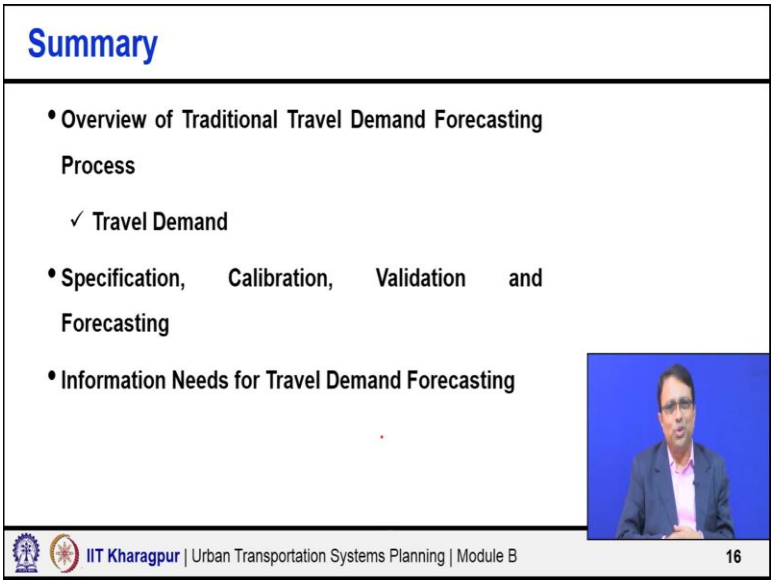
So that is what we said that significant amount of work is necessary to accumulate information and then you actually go for this one.

(Refer Slide Time: 37:29)



So all together, you know, we discussed this, next is basically what are the information needs for travel demand forecasting? We need to large number of information if we want to do travel demand forecasting. This information is primary. We can classify into four groups study area, urban activities, transportation system and travel related information. These are the four major areas we need lot of data to do the travel demand forecasting. So I will discuss in my next lecture about this information needs under these four heads.

(Refer Slide Time: 38:09)



So what we discussed in this lecture, we again give you an overview of the travel demand forecasting process, how we develop travel demand models using four stage planning process in the Base year. How we apply it in the Horizon year? Initially Base year and Horizon year or both I have discussed. Then have shown in details also how in Horizon year the captive part and the choice rider spot how they are taken care.

Then also we told you how this terminology we made you familiar with this terminology, what is model specification? What is model calibration? How model validation is done in the context of four stage planning process and then what is forecasting? And we just told you that what are the basic areas you need the data for doing the travel demand forecasting? With this we close and we continue in the next lecture.