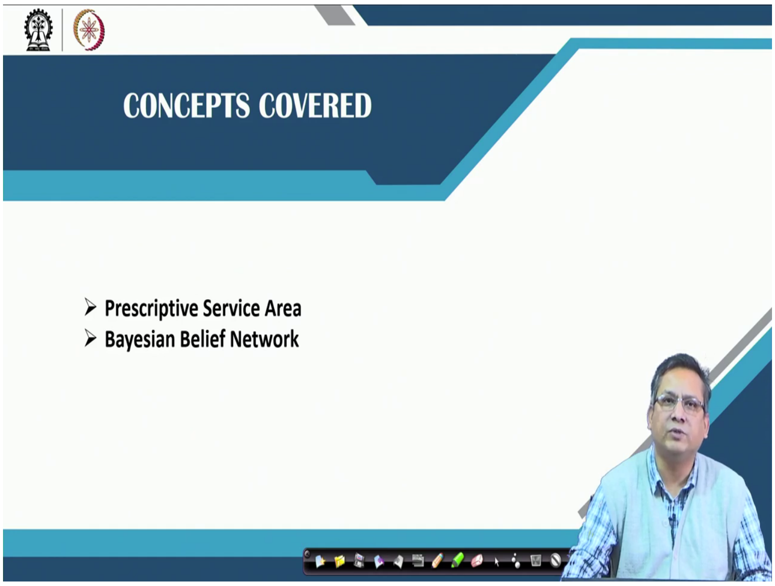
Geo Spatial Analysis in Urban Planning
Prof. Saikat Kumar Paul
Department of Architecture and Regional Planning
Indian Institute of Technology, Kharagpur

Module - 04
Modeling Spatial Relationship
Lecture - 16
Service or Trade Area Analysis in an Urban Area (Contd.)

Welcome back dear students. We are in the last week now and we shall continue with the Service and Trade Area Analysis, that we were I mean working on in the last lecture that is the lecture 15.

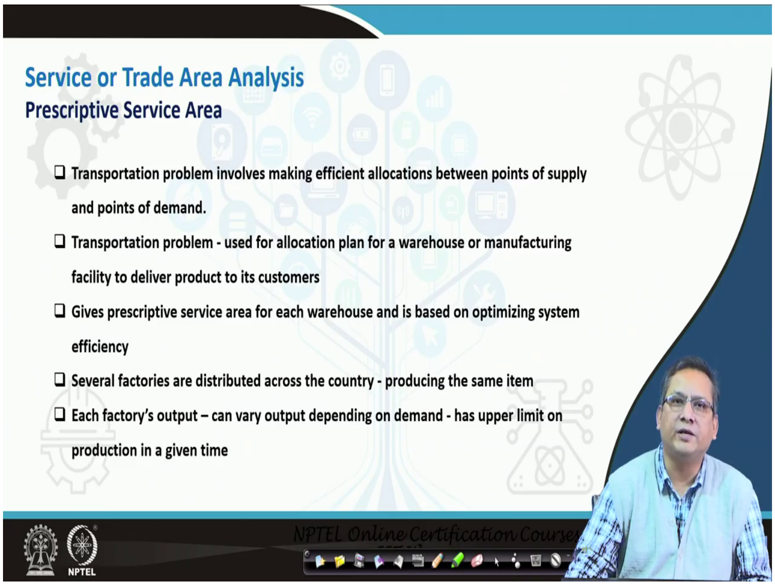
(Refer Slide Time: 00:30)



So, in this particular lecture, we are going to cover two aspects. So, first is the prescriptive service area identifying the prescriptive service area. So, in this particular approach, we use a linear programming approach to identify the prescriptive service area. In which there are multiple factories, production units, there could be transportation issues, there could be issues related to cost of production. So, this is more involved and complicated than the descriptive model that we had talked about in our earlier class.

So, then we are also going to look into to Bayesian belief network, wherein we use this Bayesian belief network, based on I mean probabilities. To see how I mean we can identify areas and we can compare it with different kind of scenarios and do a sensitivity analysis. So, we will look an application of the Bayesian belief network and we shall look into the prescriptive service area.

(Refer Slide Time: 01:38)



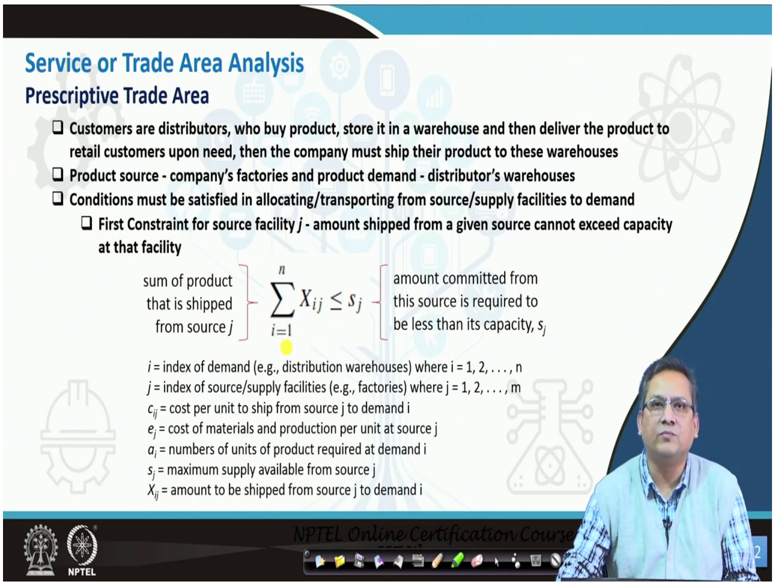
Now, when we are talking about a prescriptive service area, it is mostly a transportation kind of a problem. That is it involves decision making regarding efficient allocations between the points where a product might be in demand a from it is with respect to the supply points. So, these points may not be single locations, but may be multiple locations.

Now, your such problems could be like location of a manufacturing facility, some commodity which is being developed for I mean the consumers. So, we can have manufacturing facility or we would be we should be able to locate a warehouse optimally. So, this is a transportation problem by it is measure. So, it gives the prescriptive service area of the warehouses and we do a optimization based on the system efficiency of the entire different variables and the condition analytics.

So, I mean let us say that, we may have different types of manufacturing units or say factories located across different parts, but with the limitation is that all these industries or factories they produce the same type of item. So, let us start with this kind of an assumption. So, these outputs of these industries can also vary based on how much demand is there in a particular zone? And, these the production or the output can also have a threshold, that is it could have a upper limit of that this factory could produce so, many number of items within a given time.

So, we would be having some thresholds regarding the production and I mean some condition analytics, which needs to be encoded which needs to be looked into.

(Refer Slide Time: 03:52)



Now, we have talked about the production if we look into the consumption aspect, we would have consumers and to take this product to the consumers, what the company would do is it will store it in some warehouses. So, that the I mean product could be distributed and delivered to the customers through some retail establishments. So, the problem is that the products has to be delivered to these warehouses.

Now, we have three aspects in this particular problem, that is we are talking about the manufacturing facilities, we are talking about the demand and the location of the demand as well as the location of the factories. And, we are also talking about the warehouses from these where these products are going to be shipped to the customers.

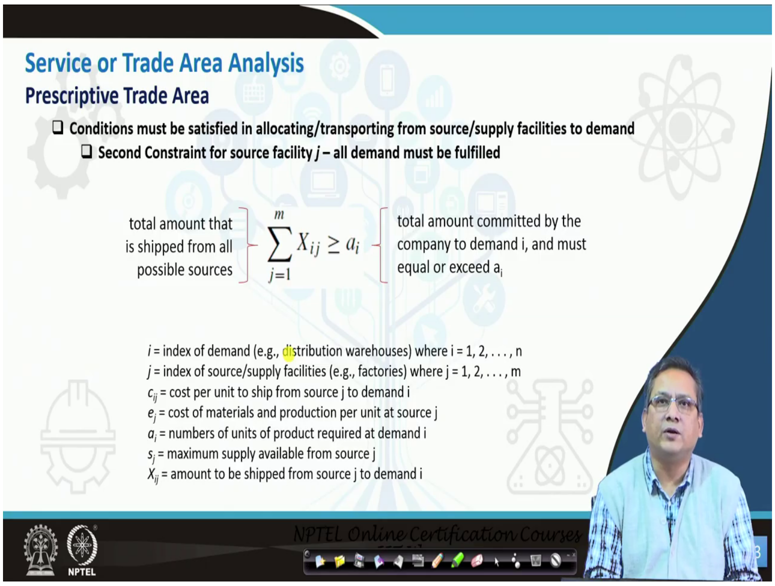
So, there would be distributors or retailers who would be having their warehouses throughout the country and these products would go into each of those warehouses and we are also talked about a threshold of production. So, certain conditions has to be satisfied, when we are I mean transporting, or assigning, or allocating, from different source facilities or that is your warehouses to the demand zones.

So, talking about the different conditions it could be written as a constraint. So, the first constraint that comes under the picture is the constraint for the source facility. Let us term this source facility and write it I mean by a notation j. So, this would be the amount shipped, but it cannot exceed the existing capacity of a facility. I mean that each of those warehouses would have a storage capability.

So, I mean we cannot ship an amount which should exceed the storage capacity of that facility. So, that becomes our first constraint. So, we can write it as sigma of X i j is less than or equal to S j, where on the left hand side your X i j is the amount of goods or say a produce, that is to be shipped from say source j, we had earlier said that I mean the first constraint is the source facility j.

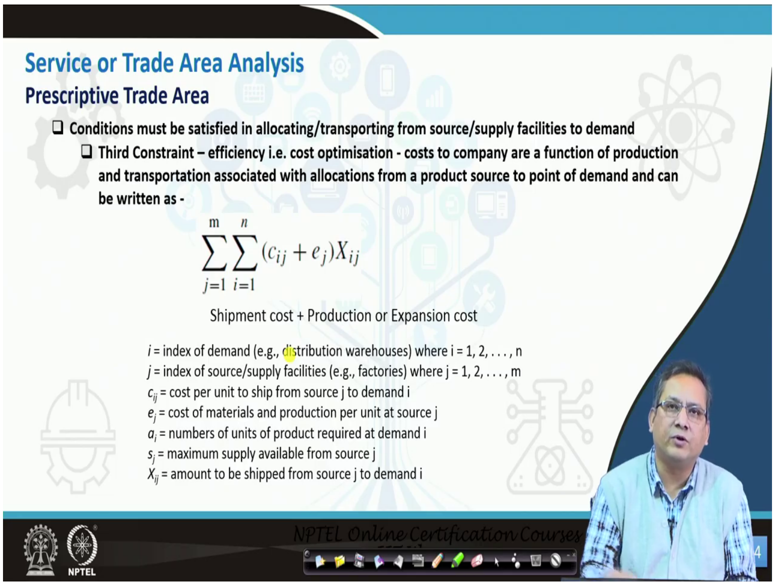
So, from source j to say demand i. So, we have number of facilities where it needs to be shipped, that is i ranging from 1 to n. So, the left hand side is a summation of the product that is to be shipped from source j. And, which we said has always to be less than the I mean storage capacity at j. So, I mean it has to be less than equal to s j that is the maximum supply available from source and which has to be less than the capacity at j.

(Refer Slide Time: 07:17)



Now, let us talk about the second constraint. So, the second constraint for the source facility is that, that whenever you have a demand, whatever demand it may be it has to be met by the source facility and the production. So, your X i j that is the total amount of commodity to be shipped from source j to say location i, I mean has always should always be greater than a, a i, that is the amount committed to the or by the company to the demand i. So, it should either be equal or it should exceed that particular demand.

(Refer Slide Time: 08:05)



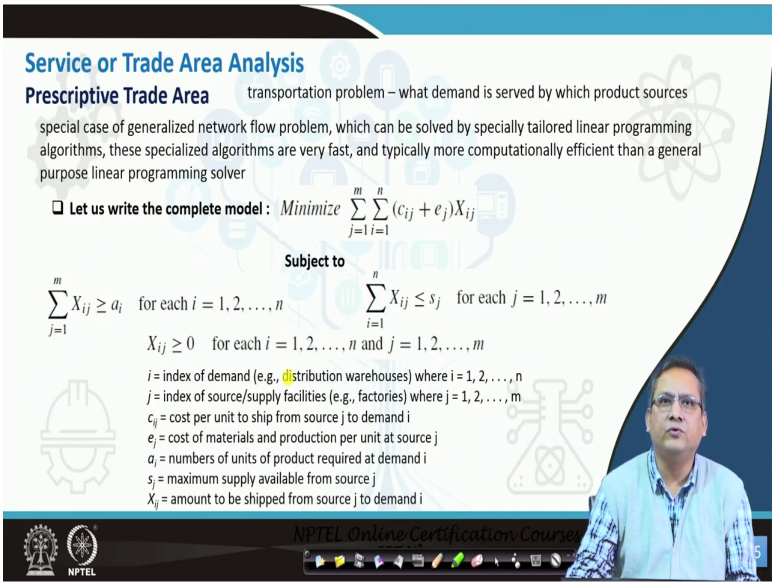
So, now we had already talked about two constraint. So, the third constraint is related to the cost and you can also, I mean have a time as a constraint, when you have perishable goods, when you are to transport perishable goods, then you can have a time as a constraint.

So, if you are designing for say location of say fire fighting services or policing. So, those I mean kind of problems would bring about different kinds of your thresholds of constraints ok. Now, we in this particular example we are talking about the cost as the third constraint. So, in those different kind of problems, you would have to identify what are the different optimization issues in those problems.

So, I mean we know that the cost of a particular good, that is I mean what the company, I mean prices the goods is a function of the transportation as well as function of the input materials the produce materials. So, I mean we have to find out that I mean this can be summed up it can be written as shipment cost plus the production or it could be a expansion cost of the unit.

Now, these two are to be I mean multiplied or it is a product of the cost plus the shipment cost into the amount of goods that are to be shipped from j to location i. And, different locations from i 1 to i n and we may have similarly different production centers from j equals to 1 or 1 to m, I mean m number production facilities.

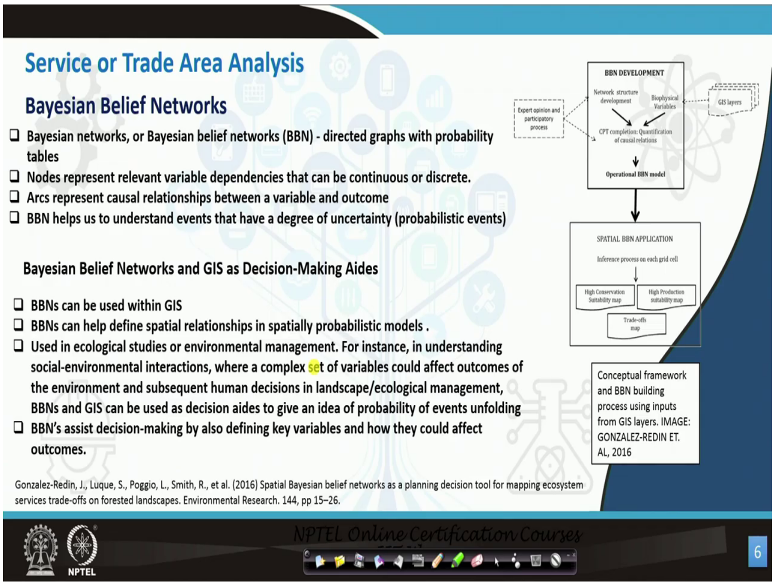
(Refer Slide Time: 10:11)



So, if we talk about this transportation problem. So, we can generalize it, we can write the complete model now. So, what we do is we have a function which is to be minimize, that is sigma j i to m, sigma i to n, summation of your cost function plus production and plus transportation function to the product of the number of units that is produced. So, this is subject to the earlier constraints that we had talked about that is X i j summation of X i j from j 1 to n has to be greater than a i and second is X i j has to be less than s j.

So, we have talked about these limitations. So, in this case X i j has to be always greater than equal to 0 and I mean your we had talked about i varying from 1 to n and j varying from 1 to m.

(Refer Slide Time: 11:15)



So, this problem can be solved using your I mean you can solve it using linear programming. So, there are specially tailored linear programming algorithms, which can solve this. So, these specialized algorithms are generally fast and more computationally efficient than the general purpose linear programming solvers. So, such solvers, I mean you can come across such solvers, there are solvers like Lindo and you can could this kind of solvers in your using Java Script and run it in Arcgis or Qgis software.

So, we can develop this kind of I mean optimization techniques using your linear programming approaches for prescriptive trade area identification. Now, let us talk about the Bayesian belief network. Now, the Bayesian belief network is a directed graph, I mean your nodes are I mean linked with your edges these are directed graphs and a directionally in nature and it has probability tables associated with these edges.

So, these nodes they represent of the relevant dependencies of variables, it could be these dependencies could be continuous or these dependencies could be discrete. Now, when we are talking about a discrete or a continuous variable, we have seen in our earlier lectures with that we had discussed in week 2, that how we can represent a continuous surface in GIS? So, we can use a raster a modeling approach, wherein we can create a raster surface, wherein we can find out these continuous dependencies or we can create polygons or vector entities for discrete variables.

So, the edges or the arcs they generally in this BBN that is a Bayesian Belief Network represent kind of a casual relationship between the variable and the I mean result that is a outcome. So, it helps us I mean to understand the probability of an event and we have when we have a given degree of uncertainty. So, it gives us the probability of occurrence of an event.

So, BBN can be used in GIS and it can aid in modeling. So, we can define the spatial relationship, using different probabilistic approaches or models. It is used for different types of studies primarily ecological studies or environmental management studies, we shall look into one such example, wherein we would be where BBN Bayesian Belief Network has been used, for assessing the flood damage assessment or flood vulnerability assessment.

So, I mean we can also workout the social environmental interaction and understand, how the different complex variables have an impact on the outcome. And, this we can use it as a management tool in decision making related to landscape and specifically ecological management. So, in a urban I mean area, we can take up decisions regarding say ecological assessment and it is management and we can also do land use and landscape assessment.

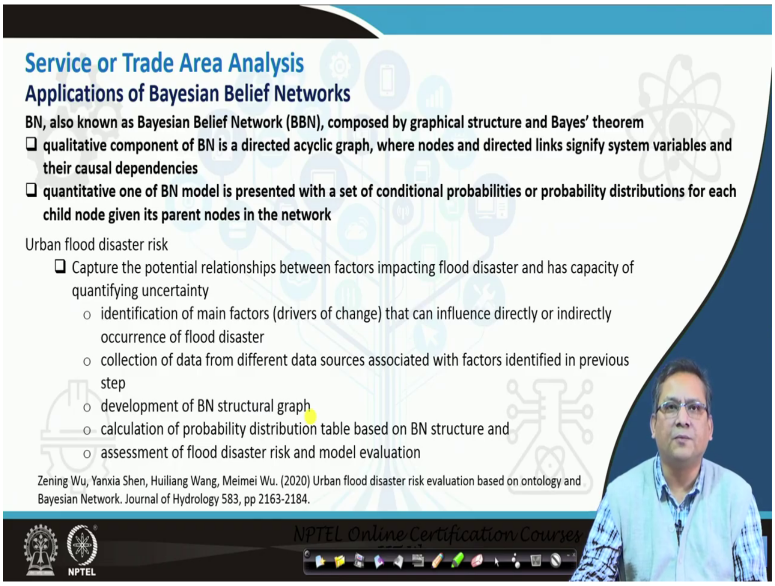
So, we can access by defining key variables and we can see how it impacts the outcome? So, in this particular flowchart, you can see that for the BBN development, what we do is we first develop the network structure, wherein we have different types of variables, which are inputs and which are GIS layers. And, then we may have expert opinions or some outputs or data from existing cause and affect relationships.

So, we do a quantification of the causal relationships. So, our different variables comes from the GIS layers and the quantification factors comes from either opinion expert opinion survey, or the I mean the causal your the models, I mean our data, that if some event happens I mean what is the probability of each factor contributing to that particular event, or alternately we can go for expert opinions service, which is a participatory process and we can do a quantification of the causal relationships.

Now, we convert this into BBN model, Bayesian belief model. So, we develop a Bayesian belief spatial application, in which this inference is processed on each and every grid cell. That is the quantification causal relationships, that we had formulated in the earlier step is used on each and every grid cell, that is in a raster framework. So, we can create different types of scenarios, different types of zones, depending on whether it is suitable, or it is unsuitable, or I mean we can also create trade of maps.

So, I mean for this you can refer to this paper which is written at the bottom this is a paper by Gonzalez Et Al this paper is title spatial Bayesian belief network as, a planning decision tool for mapping ecosystem services tradeoffs on forested landscape.

(Refer Slide Time: 17:43)



Now, we I have to look at the graphical structure and the Bayesian Theory. So, these has two components, this Bayes theorem and I mean this BBN network has two components. So, first is the qualitative component. So, which is a directed acyclic graph and where the nodes are and directed links, it signifies, the system variables and the dependents, the causal dependents, how the outcome is dependent on these particular variable.

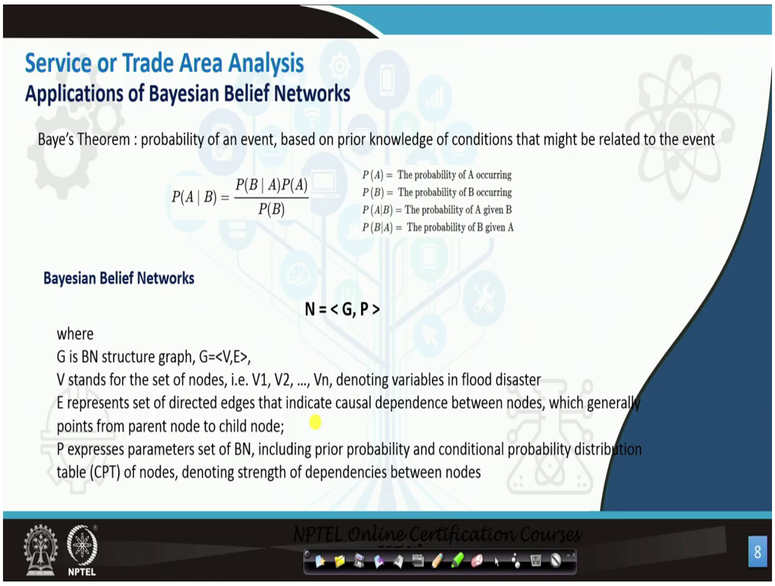
So, first is the qualitative component. Now, the second component is the quantitative one, where in this BBN model that we are talking about is presented with a set of conditional probabilities. So, these probability distributions are given in a parent and child node framework in the given network. So, these probability distributions, we had said that, we in our second step in our earlier flow flowchart, we had talked about coding these from feedback from experts or say, interviews, or we can also see the cause and impact I mean data of different variables from historical cases.

So, these can be used to identify the probability distribution for each of the child node, I mean with respect to the parent node in the network. We would be seeing this particular example in the case of urban flood, I mean what are the which are the areas which are prompt to flooding.

So, we would try to identify the factors, which would impact flood disaster and we would identify the change diverse of change which would I mean have an impact on the flood disaster. So, what we do is, to implement this we collect data from different sources. So, I mean we can identify the different sources, like in the previous step we had said that we identify the causal relationships of the input variables to the outcome, and we can identify or create GIS layers with respect to these different identified variables.

So, we develop the structural graph of the BN, wherein we have the child and the parent node and we know the transitional probabilities. So, we can use that while we are doing the calculations. The calculation of the probability I mean it is based on we create Bayesian network structure a probability distribution table. So, we shall look into it how we create that? And, finally, what we do is we can assist the flood disaster risk and we can also do a model evaluation.

(Refer Slide Time: 21:10)

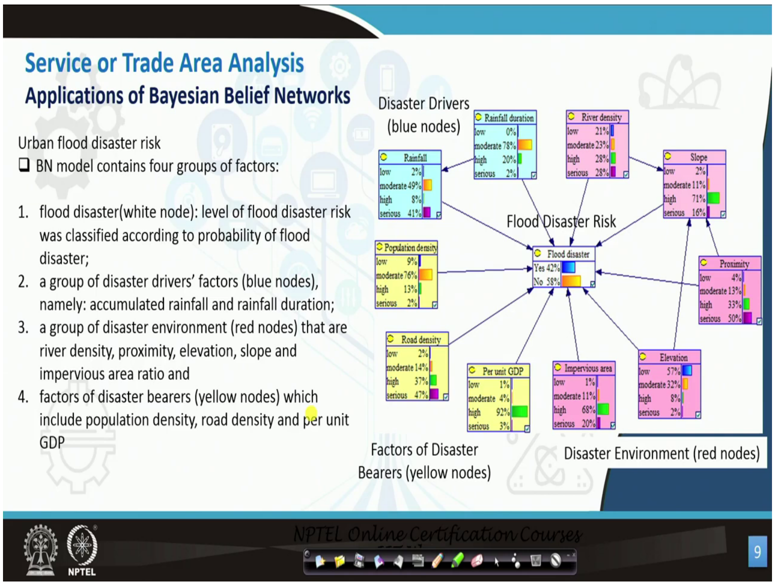


Now, this Bayesian belief network is based on the Bayes theorem. Now, Bayes theorem says, that the probability of an event, it is dependent on prior knowledge of conditions, that is related to the events or the outcome. So, we can write it as probability of A given that B is true or B exist.

So, we can find out probability of A such that B is given which is equal to probability B given A into probability of A divided by probability of B. So, we can work out the transition probabilities and we can write the Bayesian belief network, it is written as in this particular form, where in N equals G and P, where G is the Bayesian network structure graph, and G has parameters V and E, where V are the number of nodes.

I mean it denotes the variables in case of a flood disaster and your E in your the Bayesian graph G represents the directed edges, which indicates the causal dependencies between the nodes, which are generally points from parent node to the child node. Now, the P in this particular equation represents the parameters set of the Bayesian network, it is the prior probability, I mean it would include the prior probability and the conditional probability as well which is represented in a distribution table, I mean we can create a distribution table denoting the strength of dependencies between the nodes.

(Refer Slide Time: 23:15)



So, what we do is the Bayesian network model consists of four factors. So, in this particular case we the author has I mean included four group of factors. So, we have different colored nodes, I mean we will see the I mean flowchart and it will be clear to you in this course. The white node that is the flood disaster it gives us the level of flood disaster risk which is classified according to probability of the flood.

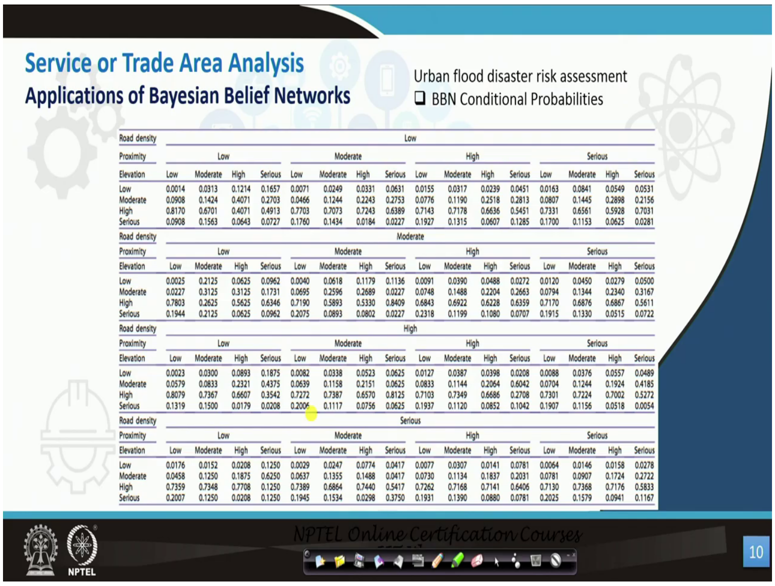
Now, the second which is which are colored as blue nodes, it gives you the accumulated rainfall and the rainfall duration. Now, there is another group which is coded as red nodes, which is the group of disaster environment thus such as a factors, which includes factors such as river density, elevation, impervious area ratio, slope proximity etcetera and the fourth one is the disaster bearer.

Now, who is the bearer? Now these are coded as yellow nodes, now, who is the bearer? It would be either the population, which is residing in that those particular areas. So, it is the population density which is going to get impacted as a result your infrastructure is also going to get impacted. So, the road density is also one of the I mean factors of disaster bearer. Next, this event also may have an impact on the economic conditions.

So, therefore, I mean the per unit GDP is also a criteria in this particular model. Now, we can see we had talked about the first one that is the flood disaster, which as which is the first which is given as the white node and it is in the center. So, you can see that if there is flood or not it would be identified based on the disaster driver, it would be based on the disaster environment and it would all also have an impact on the disaster bearer.

So, this is the entire framework of the model with the transitional probabilities that you can see. And, this is coded I mean using the equations that we had talked about in our earlier slide.

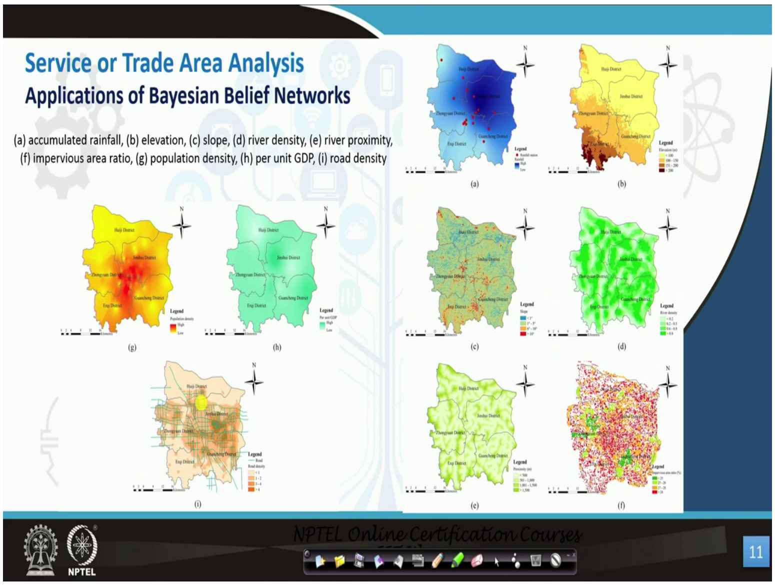
(Refer Slide Time: 25:43)



Road density	Low															
Proximity	Low				Moderate				High				Serious			
Elevation	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious
Low	0.0014	0.0313	0.1214	0.1657	0.0071	0.0249	0.0331	0.0631	0.0155	0.0317	0.0239	0.0451	0.0163	0.0841	0.0549	0.0531
Moderate	0.0908	0.1424	0.4071	0.2703	0.0466	0.1244	0.2243	0.2753	0.0776	0.1190	0.2518	0.2813	0.0807	0.1445	0.2898	0.2156
High	0.8170	0.6701	0.4071	0.4913	0.7703	0.7073	0.7243	0.6389	0.7143	0.7178	0.6636	0.5451	0.7331	0.6561	0.5928	0.7031
Serious	0.0908	0.1563	0.0643	0.0727	0.1760	0.1434	0.0184	0.0227	0.1927	0.1315	0.0607	0.1285	0.1700	0.1153	0.0625	0.0281
Road density	Moderate															
Proximity	Low				Moderate				High				Serious			
Elevation	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious
Low	0.0025	0.2125	0.0625	0.0962	0.0040	0.0618	0.1179	0.1136	0.0091	0.0390	0.0488	0.0272	0.0120	0.0450	0.0279	0.0500
Moderate	0.0227	0.3125	0.3125	0.1731	0.0695	0.2596	0.2689	0.0227	0.0748	0.1488	0.2204	0.2663	0.0794	0.1344	0.2340	0.3167
High	0.7803	0.2625	0.5625	0.6346	0.7190	0.5893	0.5330	0.8409	0.6843	0.6922	0.6228	0.6359	0.7170	0.6876	0.6867	0.5611
Serious	0.1944	0.2125	0.0625	0.0962	0.2075	0.0893	0.0802	0.0227	0.2318	0.1199	0.1080	0.0707	0.1915	0.1330	0.0515	0.0722
Road density	High															
Proximity	Low				Moderate				High				Serious			
Elevation	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious
Low	0.0023	0.0300	0.0893	0.1875	0.0082	0.0338	0.0523	0.0625	0.0127	0.0387	0.0398	0.0208	0.0088	0.0376	0.0557	0.0489
Moderate	0.0579	0.0833	0.2321	0.4375	0.0639	0.1158	0.2151	0.0625	0.0833	0.1144	0.2064	0.6042	0.0704	0.1244	0.1924	0.4185
High	0.8079	0.7367	0.6607	0.3542	0.7272	0.7387	0.6570	0.8125	0.7103	0.7349	0.6686	0.2708	0.7301	0.7224	0.7002	0.5272
Serious	0.1319	0.1500	0.0179	0.0208	0.2006	0.1117	0.0756	0.0625	0.1937	0.1120	0.0852	0.1042	0.1907	0.1156	0.0518	0.0054
Road density	Serious															
Proximity	Low				Moderate				High				Serious			
Elevation	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious	Low	Moderate	High	Serious
Low	0.0176	0.0152	0.0208	0.1250	0.0029	0.0247	0.0774	0.0417	0.0077	0.0307	0.0141	0.0781	0.0064	0.0146	0.0158	0.0278
Moderate	0.0458	0.1250	0.1875	0.6250	0.0637	0.1355	0.1488	0.0417	0.0730	0.1134	0.1837	0.2031	0.0781	0.0907	0.1724	0.2722
High	0.7359	0.7348	0.7708	0.1250	0.7389	0.6864	0.7440	0.5417	0.7262	0.7168	0.7141	0.6406	0.7130	0.7368	0.7176	0.5833
Serious	0.2007	0.1250	0.0208	0.1250	0.1945	0.1534	0.0298	0.3750	0.1931	0.1390	0.0880	0.0781	0.2025	0.1579	0.0941	0.1167

So, we have the conditional probabilities in this given table. So, this is for the different variables like proximity, road density, proximity, elevation. So, for that the different conditional probabilities have been worked out using the equation that we have talked about in our earlier slide.

(Refer Slide Time: 26:16)

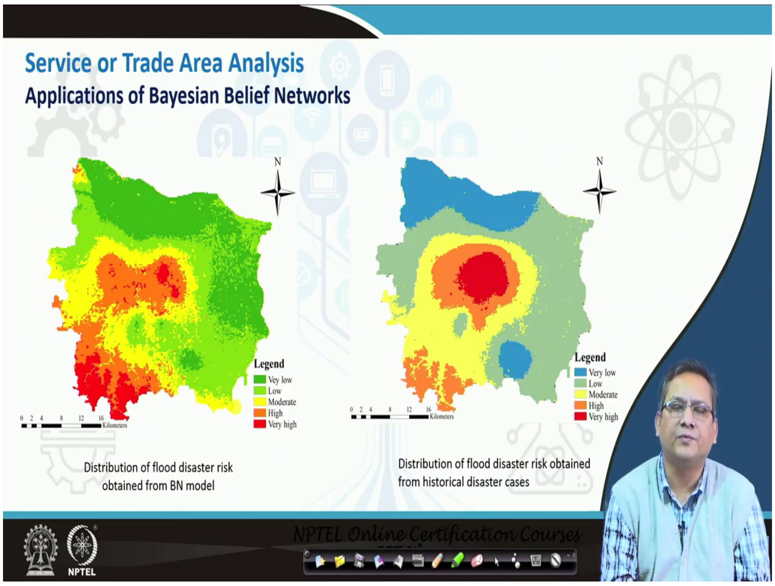


Now, we can apply this Bayesian network, I mean in the GIS mode. So, for that what the authors have done is they have the input parameters like, we said in our first flowchart we had the GIS input layers. So, these are basically the input GIS layers and we had also seen the conditional probability tables. So, you can see the first one that is the a which is the accumulated rainfall in this particular part.

So, the second input is the elevation of this particular area which is given as different colors. So, you can see from the legend what is the I mean elevation in this particular region. Now, the third one is the slope as given by degrees in degrees slope, the fourth one is the density of the river and we have the river proximity and the impervious area ratio.

Now, next we have the other parameters which are going to get affected which is the population density, then we had talked about the GDP, that is the economy of that particular area, which is going to get impacted. And, then we had talked about that infrastructure also might get affected due to floods, adversely affected due to floods. So, what we do is we identify the road density. Now there are different algorithms for calculating the road density. So, you can have a look at it.

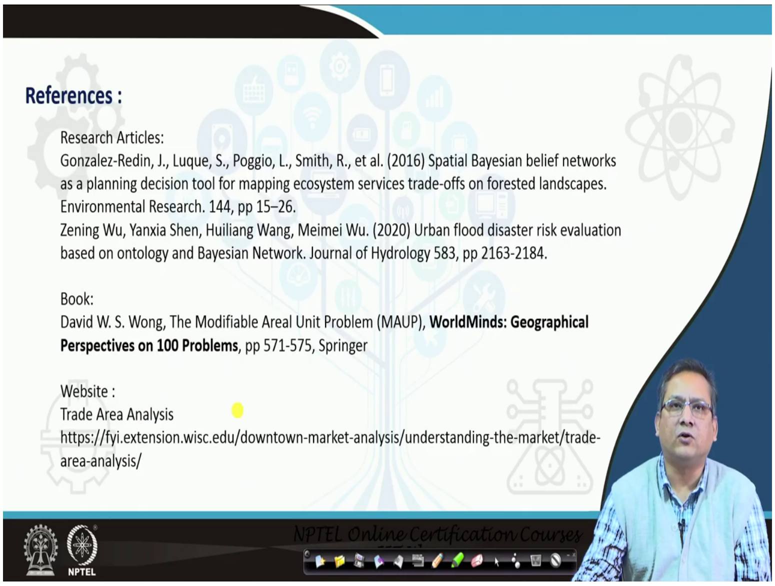
(Refer Slide Time: 28:03)



So, we can calculate the road density and finally, what we do is we aggregate we use the conditional probabilities on these particular I mean input layers. And, find out that distribution of flood disaster risk using a Bayesian belief network. So, in this you see this is the modeled outcome or simulation of the flood disaster risk map from the Bayesian network model. Now, let us compare this with the existing or the historical disaster cases in this particular area, this is the part of China.

So, the historical cases were recorded and then we can compare these two maps. So, from these two maps, you would be able to find out the degree of agreement and how sensitive your model is to developing such kind of scenarios.

(Refer Slide Time: 29:06)



So, this type of modeling framework that is your Bayesian belief network and we had talked about your I mean doing a recap, we had talked about your prescriptive service area analysis.

(Refer Slide Time: 29:09)

So, these can be used very affectively to find out, I mean the affect or the causality of a particular affect. Depending on the status of that particular area depending on the population depending on the, I mean infrastructure or it is economy. So, we can see the impact on different types of variables. So, for this particular I mean lecture I have used some references, you can look at this particular book and there is a particular chapter, which is known as modifiable area input it deals with the MAUP problem.

So, this book is titled as geographical perspective on 1000 problems, 100 problems and then we have this paper by Gonzalez and Zening, which has been used in this particular lecture. And, you can also referred to the trade area analysis website the hyperlink is given over here, which is also useful learning I mean lot of information is there in this particular website.

So, a recapitulation of what we had done today. First we had talked about a prescriptive service area in which we had used a kind of a I mean linear programming approach. So, to solve the distribution problem, move primarily a transportation problem. So, we had used the different constraints condition of constraint conditions. So, we had talked about the production cost and the transportation cost to minimize that function, using different I mean your constraints.

And, then we had talked about the Bayesian belief network, which uses the Bayes theorem and we can see the causality of a particular event given the different types of GIS inputs. And, we can create a causality I mean network in which the parent and the child edges would be given assigned a probability transition probability. And, you can find out calculate the conditional probabilities using the equations that we had discuss about.

So, using those transition probabilities you can have your input GIS layers you can also have the kind of impact that is going to be there. So, you can coded you can put it in a GIS framework and you can do an analysis of the outcome of a particular event using BBN tools. So, that is all for today.

Thanks till we meet for the next lecture.