REMOTE SENSING FOR NATURAL HAZARD STUDIES

Course Instructor: Dr. Rishikesh Bharti
Associate Professor
Department of Civil Engineering
Indian Institute of Technology Guwahati
North Guwahati, Guwahati, Assam 781 039, India
e-mail: rbharti@iitg.ac.in
Website: https://fac.iitg.ac.in/rbharti/

Lec 11a: Introduction to Geographical Information System for Data Analysis Part A

Welcome to Lecture 11. This is the continuation of the Remote Sensing for Natural Hazard Studies course. So, we have started discussing the GIS in Lecture 10; now we will continue that. So, before I proceed, I just wanted to give you a brief introduction to GIS. So, here you can see that GIS is a system designed to capture, store, manipulate, analyze, or manage datasets. Here are a few important terms. The first one is to capture, then to store, then to manipulate, analyze, manage, and present. So, when we talk about any datasets, we consider that you have collected a few data points in the field, which is basically the groundwater table. So, this is, let us say, the groundwater table. You also have information about the air quality. You also have information about precipitation; you also have information about the surface water availability. So, all this information will be stored against this particular point. So, how will the system remember, or how will you remember that these parameters are stored for location x? That is why whenever we talk about GIS, it has to be a spatial or geographic dataset.

Data to Information

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

DATA		INFORMATION	
1010010	1010011	R	S
1000101	1000101	E	E
1001101	1001110	M	N
1001111	1010011	O	S
1010100	1001001	T	I
1000101	1001110	E	N
	1000111		G

PROCESSING

That means the latitude and longitude information, which indicates the location, should be stored along with your parameters. Now that we have generated the dataset, we will store it in the database, then we will proceed to analyze and manage it, which means you can also update the database, and subsequently, you will generate a map that will display the information you have collected. So, at a time, you can generate one parameter of information in terms of the map. So, those will be called thematic maps. That means this particular map has a theme; it can be rainfall. So, we will call it a rainfall map. If it is air quality, then we will call it an air quality map. If it is surface water availability, we will call it a surface water availability map. So, in this way, we continue with this geographic information system.

So, here the GIS allows you to integrate many datasets together. So, today I will show you an example where you will easily understand how we are integrating all the information collected by two or three people, or maybe several groups across the globe. So, those things will be discussed today. Now, using this overlaying of the information on top of another will enable you to visualize the question, which means you will be able to put in the query, analyze the dataset, and interpret the data, and that will help you understand the relationship, patterns, and trends, and this will be very easy once you have built the database in a GIS environment, So, GIS is an organized collection of computer hardware, software, geographic data, and a group of people who are involved in capturing the data storing it and then updating the database then analyzing and displaying the information. So, we will continue this with data versus information because we need to understand the geographic information system.

So, let's go step by step. So, we will start with the data that has been collected using remote sensing, the conventional method, or from a field visit. You have collected some information. So, whether that data is information or both is different. So, they are significantly different; data is of little use unless it is transformed into information. Information is an answer to a query based on raw data. So, I have one example that will help you understand how the data is transformed into information. We transform data into information through the use of information systems. Now you have a database, so we will call it an information system; you put in the query, and then you will get the answer. So, the simplest example will be these values. So, it will be difficult to analyze what this means. But once you provide this to the system, it will give you letters. So, together they make remote sensing. But it will be very difficult for us to remember these values, as our system understands only the numbers correctly. So, this is data, and this is information. Now, once we have the information. We store it in the database, and then if we have any queries that are put towards this information system, an answer will subsequently be generated. So, this is called an information system. Now, here are the examples of information systems. So, we have all visited the library, and we never go rack by rack to search for a book. What do we do? We refer to the catalog. And those catalogs will help

you find which book is where, and then you can directly visit that particular section and the rack, and then you will take out the book. Similarly, the phone book has been arranged in a very easy way with the alphabet. So, if you want to identify or know the phone numbers of x, y, z people, you refer to those with the first letter of their names. Nowadays, there are different catalogs that are available. So, those will be used here when we talk about the digital information system.

Digital Information System is an integration of hardware and software to enable advanced research, communication, education, awareness etc.

Remote Sensing for Natural Hazard Studies   Dr. R. Bharti

So, a digital information system is basically the data in digital form. So, all the systems are connected to each other through the internet, and together they are helping to provide an answer to any query that anyone submits. So, here you can get the answer whenever you have a query. So, this is how we are all connected. So, we all have smartphones, and every time GPS is on, the data sharing policies are very different nowadays. So, knowingly or unknowingly, we are sharing our information, and that is helping others to get information about a particular problem. So, if you want to reach from A to B, since you have shared your information, it took you 10 minutes to reach from A to B. So, the next person, when he or she searches for that answer, will see that it will take 10 minutes to reach A from B, considering certain traffic, road conditions, and weather conditions. So, I hope both this information system and the digital information system are understood.

Now, when we talk about Geographic Information Systems, the spatial or geographic ata is involved. So, here you have a major section which is called data acquisition, then you have database management, and then you have qualitative and quantitative analysis. So, here, this analysis, the sky is the limit; any AI, ML, or machine learning algorithm can be

used under the umbrella of the GIS domain, when we talk about a particular problem, let us discuss a specific hazard; let us say we want to generate a landslide susceptibility map. How will we perceive this problem and how we will start working with GIS correctly. So, using remote sensing GIS data, field data, conventional information, and historical data, how will you be able to use them together? So, the first step to see this particular problem is to define your issue.

So, here we have defined that we want to have the landslide susceptibility map, then define the GIS criteria, and then build or import the database. Then, GIS analysis here can be anything: AI, ML, machine learning, or the simple regression analysis that is also possible. Then you are going with the output. And once the output is generated, this will provide you with the option to decide. So, here is a decision support system: whether this result is good, whether it is helping you or not; if not, then you need to modify your problem or you have to modify your GIS criteria. So, once you go through this entire channel, this is called the typical structure of GIS.
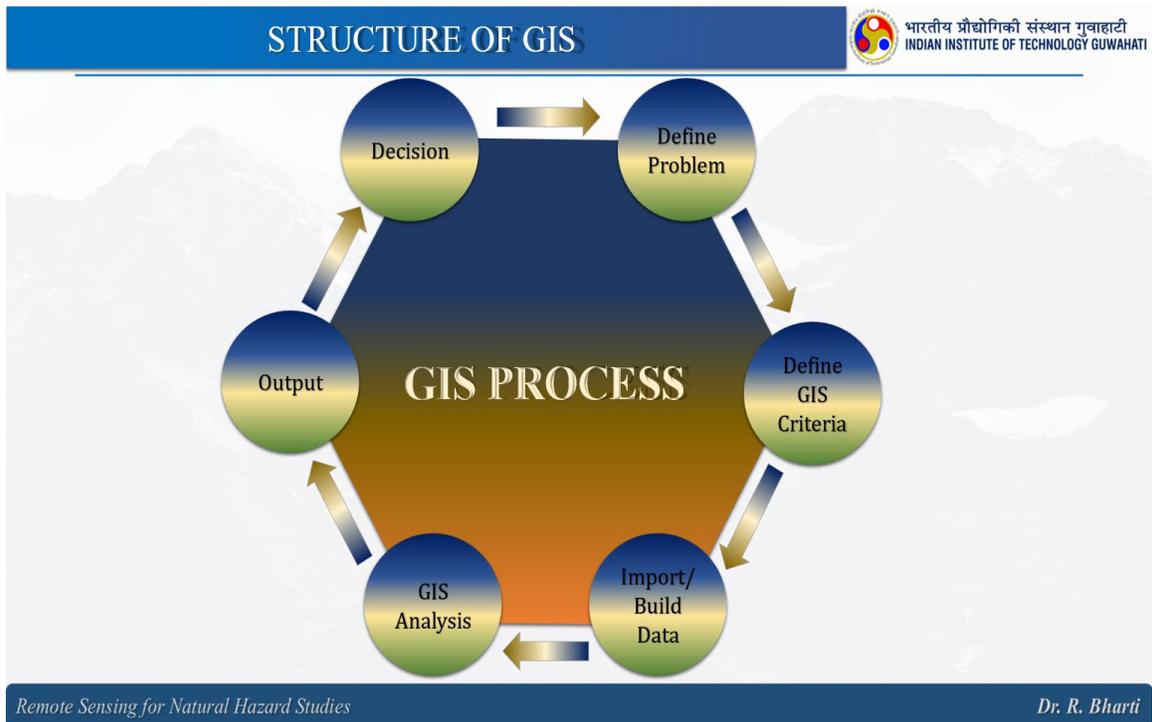
So, this is what I was talking about: the integration of the digital information.

Now, let us first group group 1. They have generated capital information for a particular country; all this capital is the center of the capital that has been generated. So, they have generated the point map for the capitals. Now, the next group is group 2. They have generated the state boundary. So, this is a polygon for the state boundary. Now the third group has generated the lake information that is present in a particular country. Now, group 4 has generated the roadmap, and group 5 has generated the information about

rivers. Now, if you look individually, you will be able to tell how many capitals there are for a specific country. Here, you can say that this is the boundary of a particular state. Here, you can say that these are the lakes present in this particular country. Here, you can say that this is the road network for a particular country. Here is the river system, or fluvial system, of a particular country; all of them are individually analyzed, but this GIS gives you the option to integrate the entire data set together. And then this map will be something like that. Here you can see that one state, let us consider this as one state. So, within this, whether a river is present or not, whether the capital is located at the center or at the corner of this state, and then whether it is connected or well-connected to the road network or not.

So, that information can be easily understood if you have digital data and in the GIS environment. So, the GIS environment gives you more flexibility to analyze your data to extract the information that was not possible to analyze when you had it in separate maps. So, here you are putting everything together in one place, and then you are extracting the information. One very important output of a GIS environment is the map itself. That map can be generated where you have the north-south direction information and the scale that will help you know how far apart these two points are on this particular map. Then this legend will help you gain the information about the different colors or symbols used here correctly. So, this map is very important because many times we need a map representation for our analysis, and that is possible with GIS. So, it is an informative 2D representation of 3D space; basically, we are trying to convert this 3D information of our natural surface to 2D. So, this is a map; it can be a hard copy representation on a sheet or a digital representation on a screen. Basic map characteristics tell the user where the object is and what the object is. So, here the location is very important, and the second thing is its attribute.

Maps are equipped with a scale legend and are defined by a projection that helps in visualizing the curved surface of the Earth on the flat surface of paper or screen. So, here we will see one by one all these elements or characteristics of a map, which will help you convey the information in a true sense when you are generating a map after your analysis. So, here is an example of an analog map, where it is sketched, and all these are approximate. So, this is not on the scale. Is it clear? But when we talk about the digital data and digital maps here, the area direction will be proper because we are using a reference system to identify the position of the object. So, maps in a GIS environment are used for input data, so they can be digitized or scanned; querying maps allows for questions such as whether A is near B, whether A is far from B, and what locations are within a 10 km radius of this particular point.

So, those things can be done. Display the result of remote sensing and GIS operations once you have derived the result or generated a landslide susceptibility. So, you want to represent the information. So, you need to provide, or you need to generate an image.

Execute and support visual analyses. And remember, maps are not only the final products; they help you to present the information, but they are not the final products. This can be input to your decision support system. So, maps provide answers to questions related to the basic components of geographic data. So, the first one is geographic location, then you have the thematic attribute, and the third one is time. So, these characteristics of the map are very, very important because the location is very important. Without attributes, your maps will be meaningless; nobody will be able to understand them, and time is important, whether this particular data was generated in 2025 or in 1980.

So, without this time, your information cannot be understood. So, when we talk about the geographic location, here you see the location of this particular player, or you can consider this a lake. So, what is the position of this particular lake, and how was it generated? So, it refers to the geographic location of a place on Earth defined by latitude and longitude. Longitude denotes the east-west position of a geographic location, while latitude runs north and south between two poles. Prime green, which is the meridian assigned as the zero line for longitude, denotes this north-south position of a geographic location, also called parallels, and runs perpendicular to longitude. So, this is latitude. The Equator is assigned a zero line for latitude; all latitudes are parallel to each other. So, now based on this latitude and longitude, we are able to identify a particular position or target in the field as well as on the map. Then comes the thematic attributes. So, a thematic attribute is like generating a map using a theme. So, it can be rainfall, it can be temperature, or human characteristics like population, health, etc. It can be used to generate the map.

So, these are examples of thematic attributes. So, here is the annual mean temperature map of India, the annual mean rainfall map of India, and here is the elevation zone of India. So, here we are only talking about the attribute, but just remember the previous slide and refer to the time. If time is not mentioned here, the annual mean temperature is not constant. So, without time, this map is again meaningless. So, without time, this map is meaningless. Now, the same thing happens with the annual mean rainfall map, but the elevation zone does not change very frequently. So, here, though time is not very important, it is required because the topography is also changing with time. This is another example where you can see a particular river with reference to 1988 and how it has migrated from this yellow line to this orange line. So, how is it changing its position now in 2002 with respect to 1988? How has this changed, and from 1988 to 2014, how is it shifting? So, the whole river has shifted. So, this is a very good example of how thematic attributes work and what information you can display using thematic information.

This is another example where this time is referred to: January, February, March, April, May, and June. So, in all of this, you see the color change, and as we have learned, if this

is a false color composite, red color basically represents the presence of vegetation. So, how is it changing? So, this is the seasonal change, and here you can see the annual change: how a forest is growing. So, it is very, very important to use these characteristics of a map whenever you are generating one. So, maps offer an abstract representation of reality that is simplified, classified, and symbolized. This is how you are going to generate your image: whether it will be simplified, whether it will be classified, or whether it will be symbolized, you will be able to understand in the next few slides. So, maps are representations at scale.

"Maps are representations at scale."

Scale: Ratio between distance on the map and corresponding distance on the ground.

Scale	Distance on map	Distance on ground
1:5,000 (Large)	1 cm	50 m
1:1,00,000 (Small)	1 cm	1 km

And scale is basically the ratio between the distance on the map and the corresponding distance on the ground. So, here, 1 centimeter on the map represents 50 meters on the ground, and in another one, 1 centimeter on the map represents 1 kilometer on the ground. So, in both cases, you see the scale is 1:5000, 1:100000. So, this is very, very important whenever you are generating the maps. Now, these are the processes, or these are the characteristics of how you are going to represent your information. So, to represent the values, you can use shape; you can use size. So, these shapes can tell that this is the x parameter, this is the y parameter, this is the z parameter, and this is a. So, wherever you have this circle, you will say x is present and y is present with this triangle.

So, Similarly, these lines can also tell you the strength of the values. So, this is thicker; this is thinner. So, similarly, color hue can also be used, and here the color value can also be used. So, different colors or hues will be used to represent the values, and color intensity will help you represent the strength of the values. So, if the value is very low, it

may be light; if it is dark, that means the value is high. Different textures for different parameters, objects, or targets can also be used. Here is a concise table that states if you have shape, you can use the qualitative differences; if you have size, you can go for the quantitative differences: color hue is qualitative, color value is quantitative, color intensity is qualitative, and texture is both qualitative and quantitative. So, if you have a smaller population, and if you have a high population area, maybe a big circle or a different color with different values can also be used, So, those are the examples of how you will be generating the map and how you will be representing your values in different symbols, colors, textures, or color hues, Now, as we talked about the simplified and classified map, whether my map will be simplified or classified has to be understood. So, here is one example; this is an example of a simplified map. Here, everything is simple; all these boundaries are visible. So, this is very simple to understand; there are not many parameters here, but when you are trying to generate this symbolized and classified map. So, remember we had the classification; after classification, we are generating the classified maps, and all the classified maps have different colors or textures that represent different classes or categories. So, here are the yellow color, red color, green color, blue color, and cyan color representing different parameters. With this, I will end this lecture, and Part 2 of Lecture 11 will continue.

Thank you.