Stochastic Hydrology

Prof. P.P. Mujumdar

Department of Civil Engineering

Indian Institute of Science, Bangalore

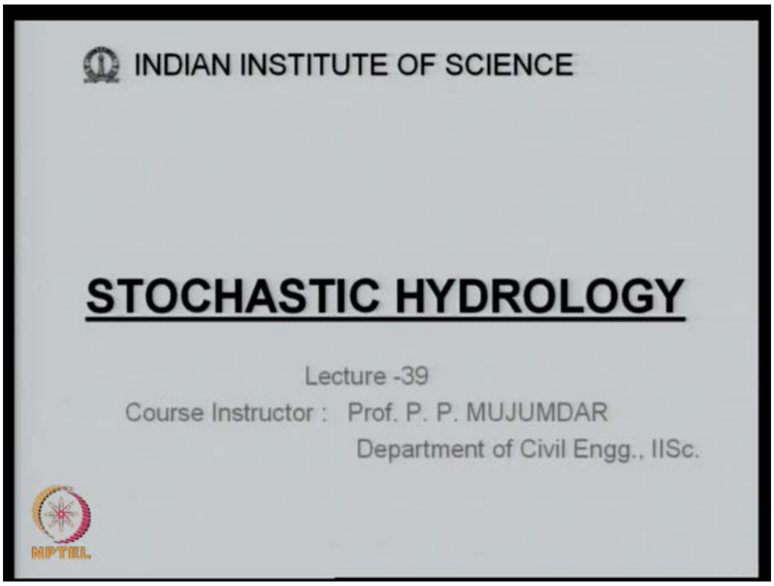
Module No. # 09

Lecture No. # 39

Recent Applications: Climate Change Impact Assessment

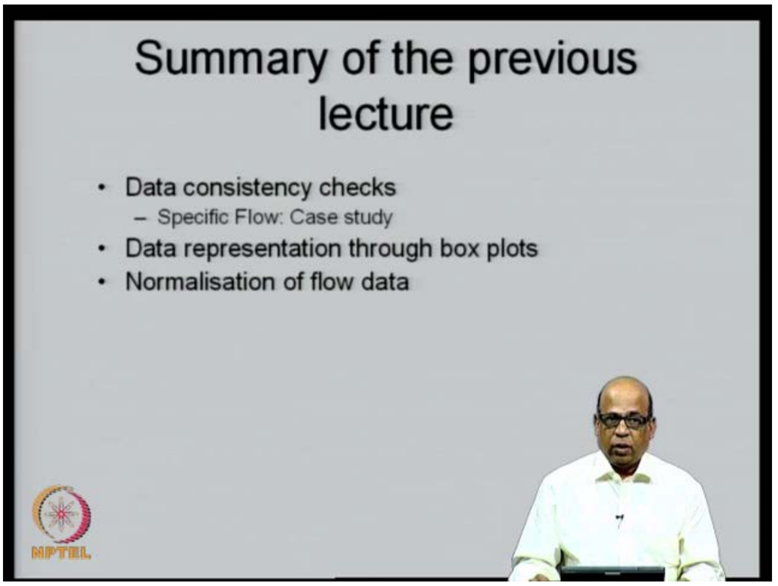
Good morning and welcome to this, the lecture number 39 of the course Stochastic Hydrology.

(Refer Slide Time: 00:22)



In the last lecture, we essentially discussed about data issues.

(Refer Slide Time: 00:32)



Specifically touching upon the data consistency checks and in the previous lecture, we had talked about the double mass curve for data consistency. And in the last lecture, we also covered the concept of specific flow. And how we use this concept to examine the consistency of data, which is in this particular method is especially useful, when we are talking about large catchments, where a number of gauges are recording the flows, specifically the flows. And then, we would like to examine whether the flows in fact are consistent with each other, as recorded from various gauges.

And we took up the case study of the Narmada basin and then, examine at various sub basins how the specific flows compare with each other. Then, I also discussed in the last lecture, the data representation through box plots. Box plots essentially will provide a good representation for the uncertainties starting with including the range, it shows the range between minimum flow and maximum flow, it also shows the median, it shows the 25 percentile, 75 percentile and sometimes we also put the mean.

Now, these box plots are useful especially, when we are comparing different time series or the variability of a particular time series across time across different time window. Let say, that you have a long time series of 100 years of flows, or you would like to examine how the variability was in the first 20 years as compared to the next 20 years as compared to the next 20 years and so on. So, these box plots are handy tools for examining such variability in the data.

Then, we also discussed in the previous lecture, how we normalize the flow data, when we have the gauge data, which are contributed essentially the flows at a particular gauge are essentially contributed from controlled flows upstream of that upstream of that particular gauge location; which means, you may have a reservoir or you may have some other control structure through which the flows are coming and then, you are recording the controlled flows at this particular location.

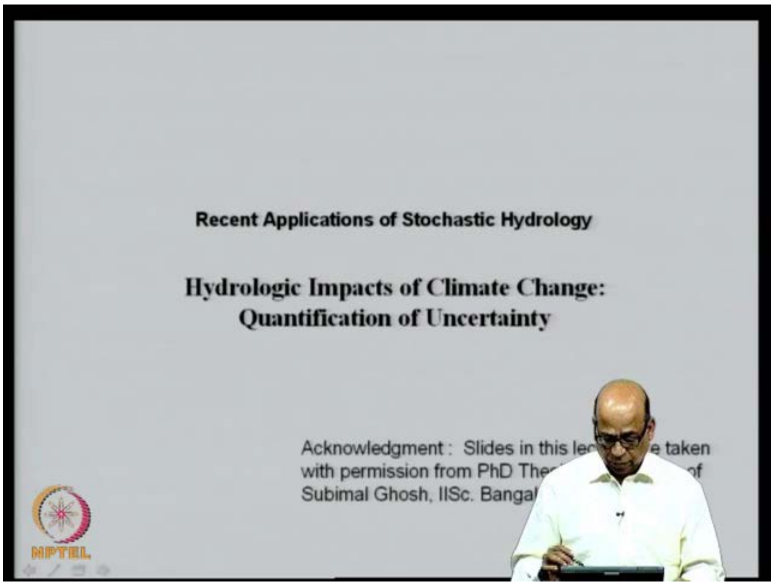
And for all analysis I have been emphasizing that, you need naturalized flows or normalized flows, which means you have to take out the effect of the anthropogenic interventions in terms of the reservoirs or in terms of at any other (( )) let say, lift irrigation schemes and so on. You may have several such anthropogenically anthropogenic interventions, structural interventions that are taken place in the stream.

And therefore, the recorded data will have signals of the controlled flows, and these signals have to be removed and then, you convert the flows as natural flows naturalized flows. And then, deal do your analysis using the techniques that we have covered in in this course on the naturalized flows. So, we have covered how to convert the flow data at a particular location into normalized flow data or naturalized flow data that essentially completes the portion that is set out for this particular course.

So, in today's lecture what I will do is, this is a special lecture as of last lecture before we summarize, I will cover a recent application of most of the techniques that we have covered in this course. And this topic is timely, because in hydrologic research, most of the most of the research that is going on is in this area. So, I will just give a broad flavor of how we use the techniques that we have covered in this course, to in this particular area, and this area is the climate change impacts on hydrology.

So, this is a special lecture, where I will just give a broad flavor I will not go into too much of details of how we do the modeling and so on. But, given that we have gone through 38 lectures so far, we will know all the methodologies that I will be discussing, so I will not go into the details of what kind of numbers we get out of methodologies and so on.

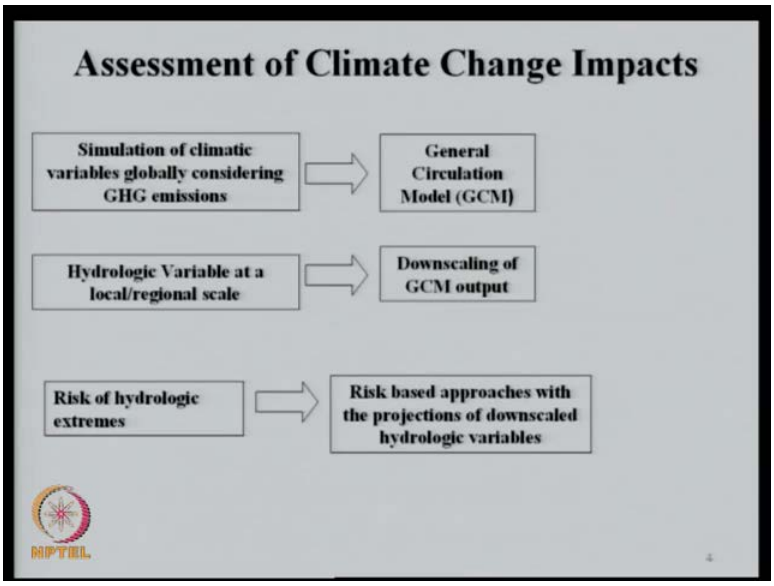
(Refer Slide Time: 05:14)



So, we will just go through the broad idea of the concepts that are involved and the modeling aspects that we do using the statistical techniques that we have learned stochastic techniques that we have learned. So, this is this lecture is on hydrologic impacts of climate change, and especially on the quantification of uncertainties.

And I have taken these slides from of my PhD students, thesis presentation, his name is Subimal Ghosh, he is already a faculty member at IIT Bombay after completing his PhD from IISc Bangalore. So, the slide set follow are essentially taken from his presentation with permission.

(Refer Slide Time: 06:02)



As I mentioned in the previous lecture, the issue of climate change impacts on hydrology is the one that deals with relating the long term changes that are happening on the climate with the changes that are likely to happen in hydrology. As I mentioned in the last class, the global average temperatures are known to be changing known to be increasing. In fact as a result of which, the sea levels are rising and the precipitation patterns across the globe are varying are changing.

All of these three prominent signals of climate change have direct implications on hydrology. They will affect the precipitation patterns, regional precipitation patterns as the result of which, the hydrologic extremes of floods and droughts are likely to change, the magnitudes are likely to change, magnitudes of the floods, the frequency of the floods and vulnerability of various regions due to floods, all of these are likely to be affected.

The water availability in a river basin is likely to be affected, the evapotranspirative demands of the crops are likely to be affected; and in countries like India, where primarily the water resources systems are operated for irrigation irrigated agriculture, these will have direct implications. And the rise in sea levels will have implications on the costal indentation, coastal flooding as well as, on the salinity ingrates into the ground water.

So, the climate change impacts need to be studied at regional scales, taking into account the projections provided by what are called as a general circulation models or the global climate models.

So in this lecture, we will see how we do that and also, how we address uncertainties associated with our projections. Because, standing at this point in time we need to assess what is likely to happen to let say, the water availability in a particular river basin over the next 20 years, 30 years, 40y years and perhaps, 1 full century; because, our water resource systems are designed to operate designed to serve for next about 100 years. And therefore, we would we should know how this systems are likely to perform in the phase of climate change over the next so many years.

So, what we do is that, we pick up the simulation of climate variables as provided by the general circulation models. Now, the general circulation models as I mentioned in the previous class, are essentially models of climate; and these are driven now by the green house gas emissions and typically the CO 2 levels, CO 2 emissions into the atmosphere and corresponding to the various levels of CO 2, emissions that are likely to happen in the future.

We have what are called as a scenarios, that is emission scenarios for example, one of the emission scenario can be double CO 2; that means, in a particular time frame CO 2 levels will be doubled like this we have several scenarios, these are given by I P C C, the Intergovernmental Panel on Climate Change, I P C C. And for these scenarios, the climate models provide us the projections of climate variables, several climate variables.

For example, what is likely to happen to temperature, what is likely to happen to relative humidity, what is likely to happen to pressure levels let say, geo potential height, that is gravity at the straight pressure levels and so on, and what is likely to happen to sea surface temperatures, mean sea level pressure and so on. So, there are large number of climate variables for which the simulation are provided by the global climate models or the general circulation models.

Now, as hydrologist we pick up this simulation of the climate models and then, relate it with the hydrologic variables. We will be interested in, what is likely to happen to precipitation in a particular catchment or what is likely to happen to stream flow at a particular in a particular river basin, what is likely to happen to evapotranspiration in the particular river basin and so on. These variables that the hydrologist are interested in, are not well simulated by the climate models for various reasons we need not worry about those things. But, the variables at a really of interest to the hydrologist are not well simulated by the general circulation models.

In addition, the general circulation models operate on large scales by their very nature, because they have to account for the entire globe, these are circulation patterns. So, which which means, you cannot afford to just look at one particular region in isolation, everything is related; and therefore, the global climate models will generate the climate taking into account, various grids across the globe and because of that, because of the computation requirements, because of the several other requirements, they essentially talk operate at large scales. What I mean by that is? The grid size of climate models are much larger compared to the size that is required for hydrologic impact assessment.

So, because of these two major reasons, namely that the variables that are required in hydrologic applications are not well simulated by the climate model. And that the global climate models or the general circulation models provide the outputs essentially at large scales, because of these two reasons we adopt what is called as the downscaling.

So, we downscale from the G C M's to the hydrologic scales, how we do this, we will study slightly, we will study in the present lecture. So, we pick up the simulations of provided by the general circulation models, we need the hydrologic variables at a local or regional scales; and this we obtained from downscaling of the G C M output, we use this to project into the future, the hydrologic scenario. So, associated with the climate scenario, we provide a hydrologic scenario.

And then, look at the risk associated with hydrologic extremes let say, the droughts, the floods and even the risk associated with operating a particular reservoir or water resource system in a particular manner. Risk associated with hydropower being lower than a particular level, water supply being a particular level and so on.

So, the risk associated with not only hydrologic extremes, but also with the various features of the water resource systems. And we developed risk based approaches and so on. So, this is a broad idea of what we do in hydrology in as much as climate change impacts are concern. However, we will focus in the in the lecture essentially on downscaling techniques and a specific class of downscaling techniques called as a statistical downscaling techniques, that use the transfer function approaches.

When we do all of this analysis, the topics that you have covered in this particular course and some adaptation of those topics will come in handy, in addressing the various levels of uncertainties (Refer Slide Time: 13:50). There are model uncertainties, because we are relying on the global climate models or the general circulation models, there are handfuls of such models in fact, there may be as many as 25 such models, for which the outputs are available across the globe and there are many more coming.

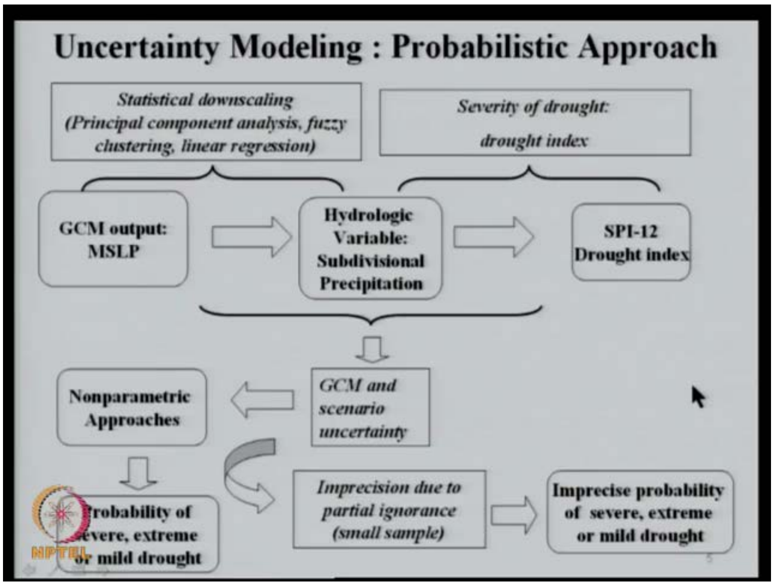
So, we have what is called as a G C M uncertainty? That is, General Circulation Model uncertainty, depending on which model you choose for the particular region that we are interested in, you may get different types of projections when you bring it down to the regional levels. There are various reasons for that, we will not in this lecture we will not worry about, why the G C M uncertainty in fact, happens.

Then, there is also a scenario uncertainty as I said scenarios are pictures of how the world is likely to evolve in in in future, now there is a enormous uncertainty associated with the scenarios. How the CO 2 levels will be in the future and because of which, how the temperatures are likely to be in the future and so on. Standing at this point in time we are trying to project it into the future and there are large numbers of uncertainties associated with scenarios.

Then within the model itself, there will be some uncertainties, because of the parameters and so on. The G C M uncertainty that I talked about, are intermodal uncertainties when you choose one G C M, it produces a certain type of projection, whereas another G C M produces another type of projection that is intermodal uncertainty. But, there are also intra-model uncertainties, which arise because of the parameters that have been put into the G C M, the way parameters are estimated, the way boundary conditions are estimated, the way initial conditions are estimated and so on. So, that leads to intra-model uncertainty.

Then as I said, from the climate models we are bringing it down to hydrologic scales on hydrologic variables and therefore, we are doing a particular type of downscaling. So, the down scaling itself produces certain uncertainty. So, there are large numbers of uncertainties when we are talking about climate change impacts on hydrology. The methods, the tools, the techniques that you have studied in this particular course will all come in handy, when we are addressing these kinds of uncertainties.

(Refer Slide Time: 16:42)



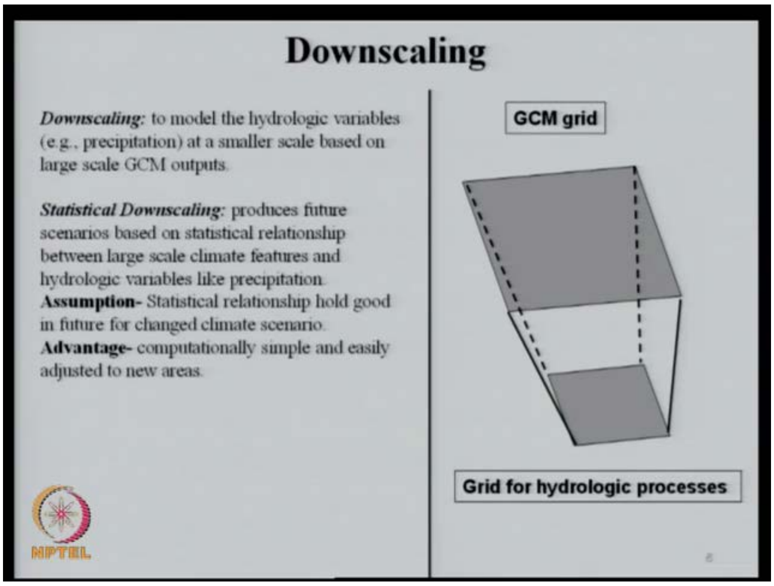
So, in the probabilistic approach I I will I will just give a broad overview of what we do here and in the particular application that I will be talking about, in this lecture that will deal with the drought index or how the drought scenario is likely to evolve in a particular river basin. So, from the G C M we do a statistical downscaling, using the principle component analysis and linear regression, we also add fuzzy clustering, but I will not cover that, because we are not covered in the course the fuzzy clustering, but typically we have studied the regression using the principle components.

So, I will just explain how these techniques are used in that. So, for doing this we need the G C M output. And the specific example that I will be covering, we will use the mean sea level pressure, so mean sea level pressure is likely to effect the precipitation directly. And mean sea level pressure is well simulated by most of the G C M's and therefore, we pickup mean sea level pressure as, what is called as a predictor of the precipitation, which is called as the predictand.

We are now developing a relationship between the predictand, which is a precipitation at a regional scale with the predictor which is a mean sea level pressure at several grid points, several G C M grid points. From this, we get the subdivisional precipitation in in this particular case, we are talking about Orissa subdivisional precipitation as I will show you presently. From this, we get what is called as the standardized precipitation index of 12 month scale and that is what we use as drought index. And we address the G C M and scenario uncertainty by fitting probabilistic probability distributions to the S P I 12.

And then, we address several levels of uncertainties, one is a data sample uncertainty itself, if your sample size is too small size is reasonably small and you are trying to fit a probability distribution to that; then the probability distribution itself will have uncertainty, that we address through imprecise probabilities. And then finally, we provide the probability distributions for the drought index into the future, so this is a broad overview.

(Refer Slide Time: 19:23)



We will go into some of the details here. So as I mentioned, the downscaling deals with bringing down from a larger scale like this, this is a G C M grid point, it until recently it used to be of the size 2.5 degree by 2.5 degree, that is 2.5 degree longitude, 2.5 degree latitude, which may be around 250 kilometers by 250 kilometers; that means, we are getting one output from the G C M for the scale of 250 kilometers by 250 kilometers; whereas on the hydrologic scale, if you want to run any hydrologic model, grid base hydrologic model like for example, variable length field trace and capacity model and so on, you will need this outputs at 15 kilometer by 15 kilometer, 12 kilometer by 12 kilometer and so on.

So, these are the grid sizes required for hydrologic processes. In fact, depending on the process that we will be interested in, these need to be even smaller of the order of 5 kilometer by 5 kilometer and so on. So, we must be able to bring down the G C M outputs to the hydrologic scales, and this we do by what is called as a downscaling.

There are two major ways of doing downscaling, one is called as the dynamic downscaling, and another is called as a statistical downscaling. The dynamic downscaling is beyond the purview of the hydrologist. In dynamic downscaling they actually develop the regional climate models taking into account the the regional climate models, take the boundary condition from the global climate models; and therefore, for a given G C M, you develop a regional climate model for a particular region.

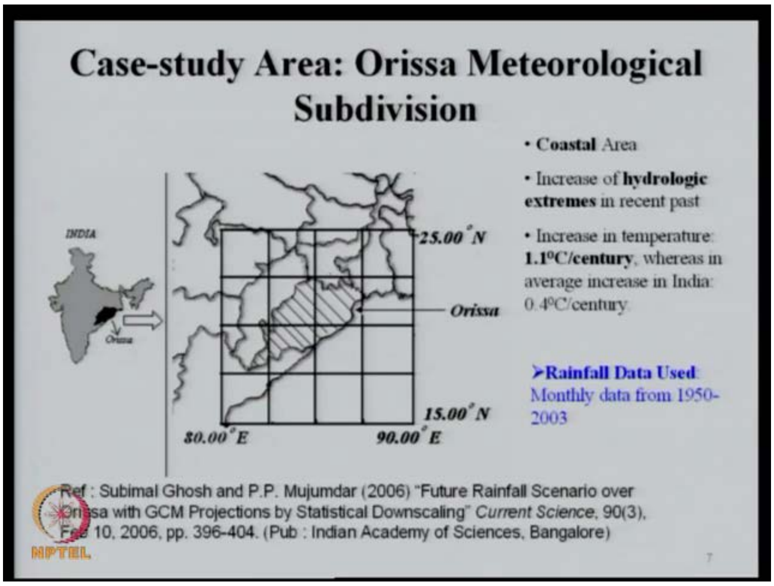
However, that is beyond the scope of hydrologist, so we adopt what is called as the statistical downscaling. The statistical downscaling produces future scenarios based on statistical relationship between the climate variables. Let say, in this particular case we are talking about mean sea level pressure, so mean sea level pressure at a particular grid point, we may have mean sea level pressure here (Refer Slide Time: 21:37), and then we are interested in precipitation at a smaller scale.

So, we develop statistical relationship between the mean sea level pressures at one or more grid points, with the precipitation at a particular location. These statistical relationships can be as simple models as simple linear regression or multiple linear regressions to very very complex models using the conditional random fields and so on. But essentially, they use the concepts of probability and statistics. Once we develop this statistical relationships based on the historical data, we hold this statistical relationship intact and then, use it for future projections.

So, the G C M's will provide the future projections, we use the same statistical relationship that is developed based on the historical data between let say, the M S L P Mean Sea Level Pressure and the precipitation. We hold this intact and then, use the M S L P as projected by the G C M into the future and then using the same statistical relationship, we project the precipitation into the future.

The main advantage of statistical downscaling is that, it is computationally simple and then any new area you can develop a new relationship for different G C M's, whereas for dynamic downscaling for every G C M, you have to develop a different regional climate model ok.

(Refer Slide Time: 23:23)



We will just look at one application here, this region that is shown is the Orissa meteorological subdivision as provided by (( )). So, we have a rainfall series available for Orissa meteorological subdivision, this **(( ))** area shows the Orissa meteorological subdivision and the location is here.

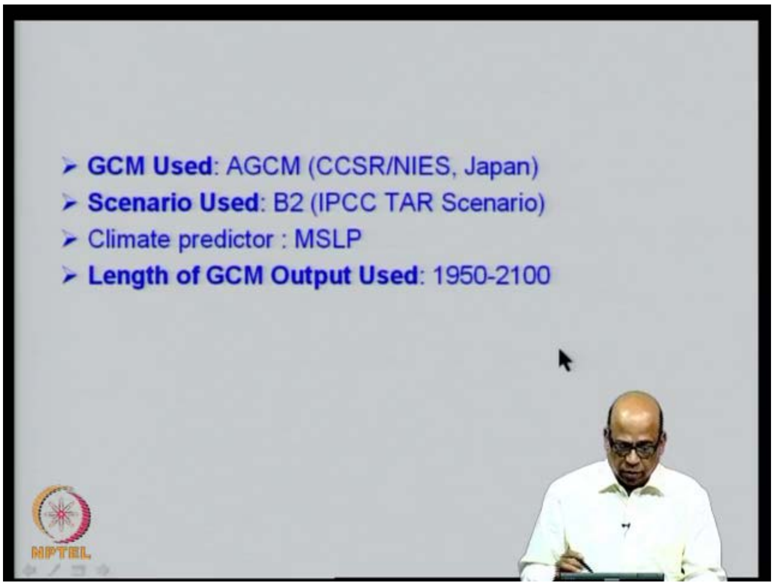
Now, this region is sensitive to climate change, because it is a first of all, it is a coastal area then, there are increase in hydrologic extremes that are evidenced in the past. It is subjected to both floods as well as, droughts frequently in the past. Then we use a rainfall data, available from 1950 to 2003, we have monthly data available. So, this time series is what is available to us, what we do is that, we superpose the G C M grid points, this is a particular G C M I will show you which G C M we are using and so on. So, you choose a G C M and then superpose on this area, the g c m.

If you have historical data at all this grid points on the climate variables that we have chosen as predictors, then we use the historical data to develop a relationship between the particular predictors at this various locations with the precipitation at this regional scales. So, we have the precipitation at this regional scale which is called as a predictand, we want to predict that particular variable; and we have the climate forcing coming from what are called as the predictors. So, the predictor in this particular case is, the mean sea level pressure, so the mean sea level pressure values we may have at various grid points like this.

In the absence of observed predictor value, data on the predictor variables we use what are called as a reanalysis data. The reanalysis data essentially provide us the closest data to the historical point, they use a large number of data sets across the globe and then, they run idealized G C M and then, provide us with the data on all the climate variables. So, we will have the reanalysis data and we also have the G C M produced data for the future.

We use the reanalysis data or the historical data to develop relationship between the predictand, which is the rainfall here and the predictors at various points and then, hold this statistical relationships (( )) intact; and then, look at the projections provided by the G C M, use the same statistical relationship and provide the projections of the precipitation, that is a principle of statistical downscaling.

(Refer Slide Time: 26:53)



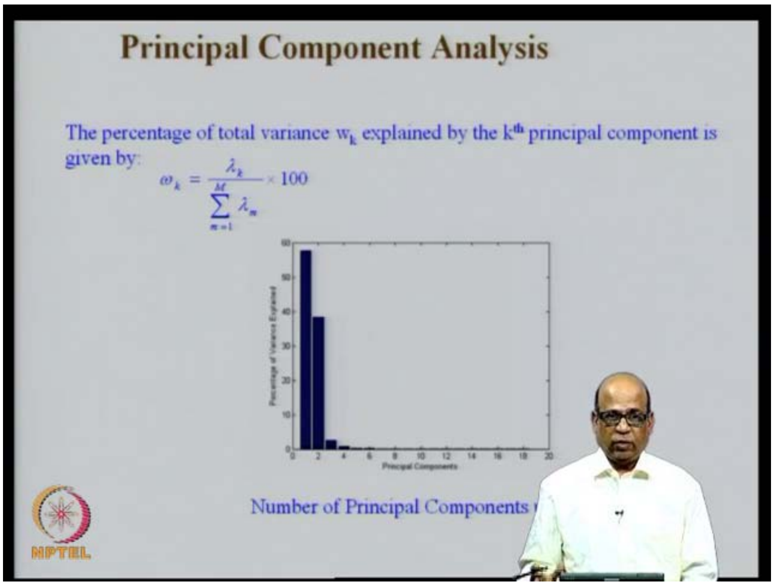
So, in this particular case, we use the G C M of C C S R, Japan, this is a A G C M; and then, we provide we used one of the scenarios just for demonstration I will use this scenario now, it is called as B 2 scenario, it is from the third assessment repowered of I P C C, Intergovernmental Panel on Climate Change.

We use the climate predictor as M S 1 P to begin with Mean Sea Level Pressure and we relate the mean sea level pressure with the precipitation at the Orissa metrological subdivision. And the G C M output are provided up to 2100, so you have certain output for calibration of your models as well as, you have the projections into the future, so we use from 1950 to 2100.

You look at this now (Refer Slide Time: 28:03), at each of these grid points you have one variable mean sea level pressure available at this locations and then; that means, sea level pressure is a time series, its available in the form of time series. So, we use the monthly data of rainfall from 1950 to 2003. So, you have the time series of the rainfall, you have the time series of the mean sea level pressure at all of these locations, you can develop a relationship between the rainfall and this predictors.

Because of the size of the problem as you can see here, if you take 4 by 4 it will be 16. And then as your predictor number of predictor's increases, you will have larger number of a larger number of variables involved in the regression, if you are using regression. Because of the size as well as, because of what we call as multicollinearity which I discussed in the multiple linear regression topic, we do the principle component analysis to remove the multicollinearity and to reduce the size. So, we use the principle component analysis to identify how much of percentage of variance, that is explained by the principle components.

(Refer Slide Time: 29:17)

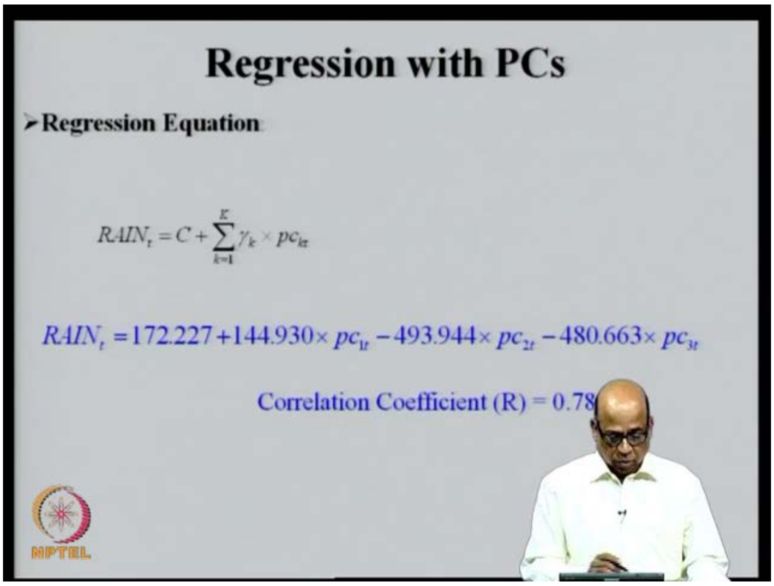


So, in this particular case, you can see first two explained almost about 95 percent, this is let say 59 and this is 37, so about 96 percentage will be explained, but we can take up to third, upto the third principle component you choose for modeling purposes; all other principle components, do not add to the explanation of variance significantly.

So, we choose three principle components. Recall your topics that we covered in the multiple linear regression using the principle components, we use the eigen eigen values and eigenvectors; and then, identify those particular principle components initial few principle components, which explain most of the variance in data; and then choose these principle components for our further analysis.

So in this particular case, we identify three principle components using these three principle components we fit a relationship between the rainfall in a particular month. This is a monthly data that we are talking about, rainfall in a particular month with the principle components.

(Refer Slide Time: 30:40)

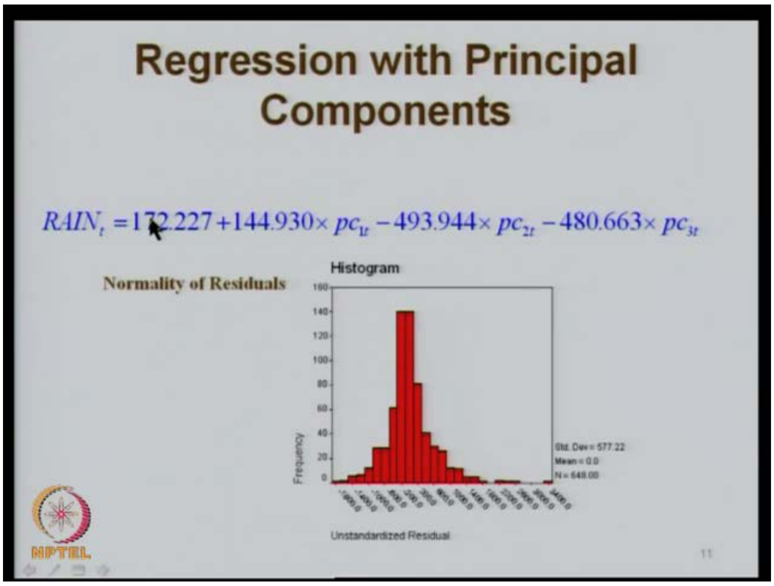


So, we write this simple expression, this is a regression equation rainfall in the month t is equal to some constant, and these are the k principle components, gamma k into p c k t, so k is equal to 1 to k. Remember here, the principle components will be different for different time periods. So, in this particular case, we have chosen three principal components p c 1, p c 2, p c 3 all of which are different for different time periods, and these are the gamma values.

So, we write the rain t as a simple regression form in a simple regression form as follows, when we do this for the training period; that means, whatever data that we have used let say, between 1950 and 2003 we fit the regression relationship and you get a R value of 0.789; which means, predicted and the observed the coefficient between the correlation between predicted and observed is 0.789, which is quite acceptable.

In fact, we do several things to improve this, let say we do clustering and then, we do fuzzy clustering and then, we add a seasonality term etcetera, all those details we will not worry about. We will say that, we start with a correlation of 0.789 between the observed and the predictor predicted. When we go into the future, now this is this equation is fit using the historical data, these pc's that are obtained are using the historical data. We use the same principle directions into the future and then, use this equation for projected values of the climate variable, which is the mean sea level pressure in this particular case.

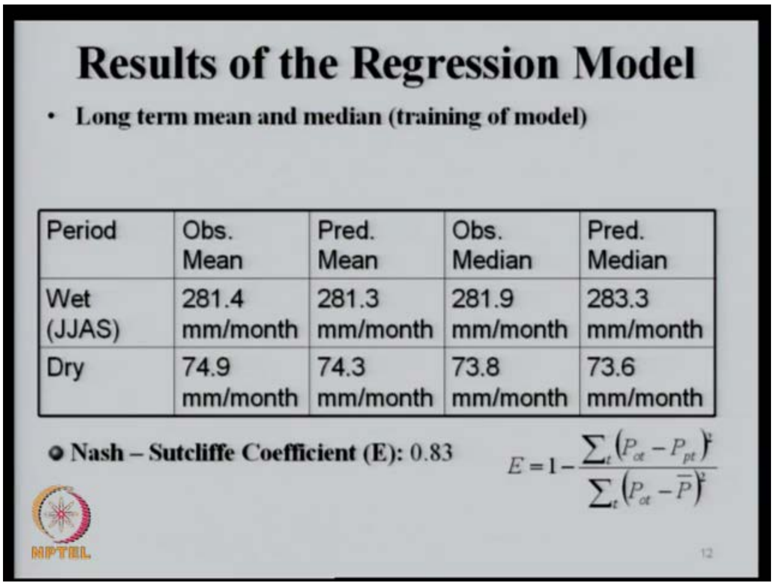
(Refer Slide Time: 32:34)



When we fit the relationship as I explained in our multiple linear regressions, the residuals have to be normal, they have to follow normal distribution and therefore, we check for the normality of residuals. What do I mean by residuals? That is, the predicted minus the observed that error, that error series is called as a residual series. So, we fit the residual series, this is a unstandardized non standardized residuals. So, as they are obtained we fit it and then see that, it more or less fits a normal distribution.

In fact, there is more rigorous test that we carry out to examine whether the residuals in fact, are normal and they are uncorrelated and so on; which I have explained in time series analysis course, all of time series analysis topics, all of those test can be carried out. So, this is just to demonstrate that they in fact fit the normal distribution, reasonably well.

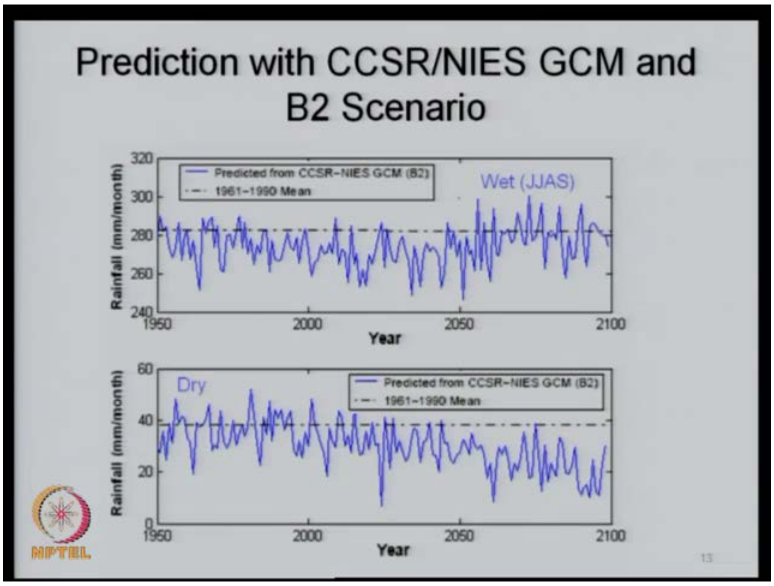
(Refer Slide Time: 33:43)



Then, we also look at this is for the training of the model; that means, when we are building the model, we look at how well the observed mean and the predicted mean agree with each other, this we do for the wet season as well as, the dry season. The wet season is our monsoon period June, July, August, September and the dry season is the remaining period, so they fit very well in fact, 281.9, 283.3 and so on.

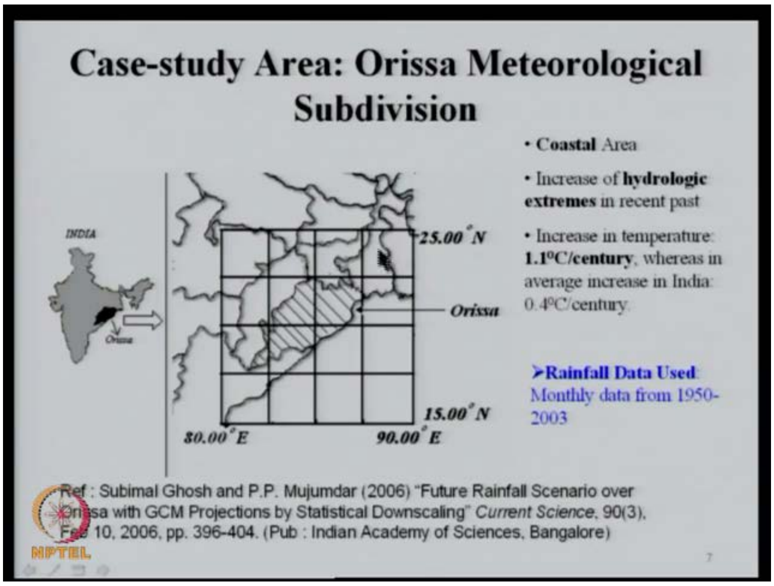
So, mean as well as median we check for the model and then, we also check with the Nash-Sutcliffe coefficient, which is generally used in our hydrologic literature, which is given by (Refer Slide Time: 34:31), this is a observed at time t, this is predicted at time t and this is the mean for the entire time period. So, this is how we calculate the Nash-Sutcliffe coefficient and this is 0.83, which is very good.

(Refer Slide Time: 34:58)



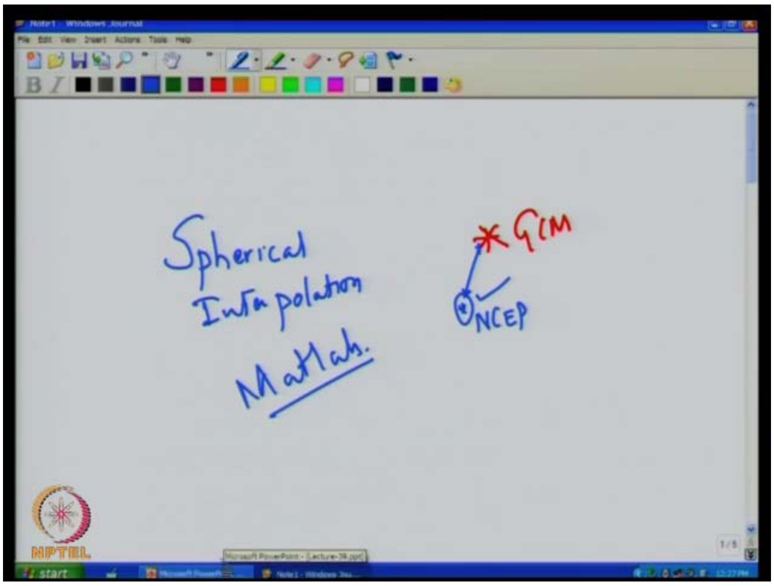
Once we are satisfied that, the regression relationship that we have built is acceptable, then what we do is? We go to this particular G C M and obtain the projections on the climate predictor. In this particular case, we have used only one predictor, which is the mean sea level pressure at those grid points look at this.

(Refer Slide Time: 35:23)



So, we are looking at the mean sea level pressure, provided by the G C M at these grid points. Now these are available for download, if you go to the G C M site you will get the simulations provided into the future for for all the grid points. So, you can extract specifying your longitude as well as, latitude you specify and then, those grid points you will get the mean sea level pressures directly from the G C M website. You use those mean sea level pressures. In fact, there is a slight (( )) here as I mentioned, we may use reanalysis data; and the reanalysis data, which is given by typically we use the NCEP and NCAR reanalysis data.

(Refer Slide Time: 36:25)



And what happens is in most cases? Let say, this is a G C M grid point, your NCEP grid point may be somewhere here, so this is NCEP grid point, this is a G C M grid point. Let say, we are using a particular G C M and the G C M grid G C M is here, grid point is here, whereas your reanalysis data is here.

And we would have developed the reanalysis we would have developed the regression relationship with the reanalysis data and therefore, when we use the projections this projected variable or the projected simulations must be brought to this grid point that is the NCEP grid point. So, we need to interpolate the values that are provided by the G C M at particular locations to the NCEP grid points, and this we do typically by what is called as a spherical interpolation.

So, we use spherical interpolation for interpolating between G C M grid points and the NCEP grid points, this is just a broad idea, but there are lots of improvements that are available that are possible in what I just told. And spherical interpolation is readily doable using matlab programs. So, in matlab you have a routine to carryout spherical interpolation you can use that toolbox and then in fact, there is a function for spherical interpolation you can use that and then, do the spherical interpolation. But what is important to understand here is that, you are developing these kinds of relationships with respect to the reanalysis data.

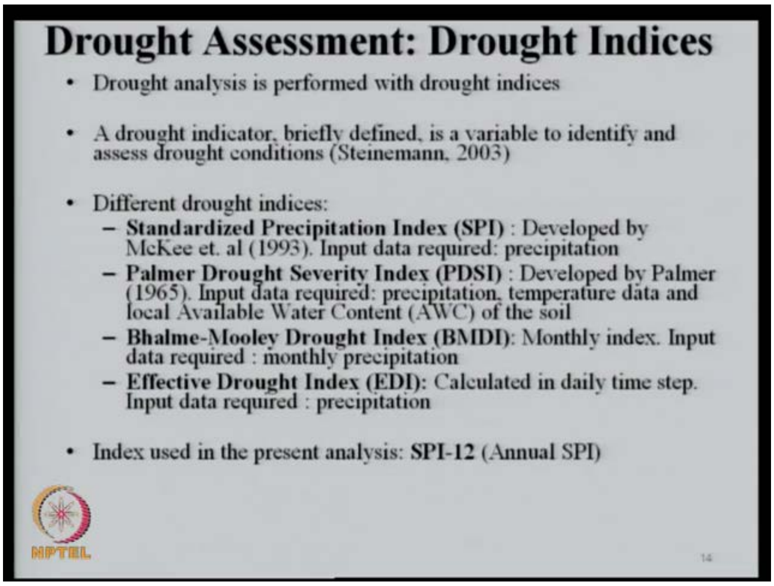
And the reanalysis grid points for example, NCEP grid point, NCEP is National Center for Environmental Prediction U S A, those grid points and the G C M grid points may not tally and therefore, you have to do an interpolation. So, when we do this using the particular scenario that we are interested in as I said, we are using B 2 scenario here; and with one particular model this is a C C S R N I E S model, this is a Japanese model, we use this for the particular scenario; and then, project it into the future upto 2003 or something we use it for building the model and then, the remaining part we use it for projecting and this is a long term mean.

So, you may see certain trends that, the precipitation mean may be falling up to certain point and then rising beyond a certain point as in time and so on. So, we see certain patterns as we project it into the future, this is for the wet scenario, this is for the dry scenario, dry scenario may be falling down like this.

However, the point that is important and the point, where the uncertainty is start arising is as you change this model and as you change this scenario, you will get different pictures all together. So, instead of this model, if you use another model, you get a different picture altogether, instead of this scenario you choose another scenario let say instead of B 2, you may choose A 1 then you may get a different scenario altogether. And that is what causes uncertainty, because you are interested in obtaining the hydrologic scenario.

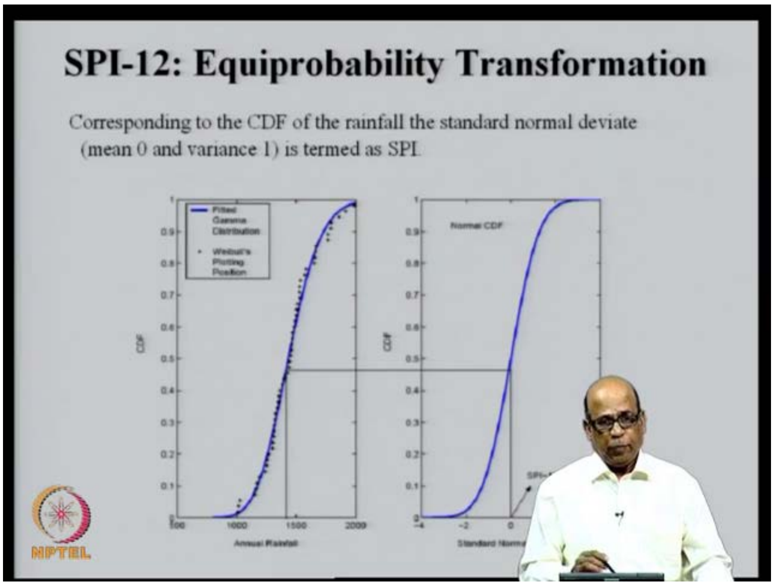
For example, you are interested in seeing how the precipitation is likely to change in future. And this uncertainty needs to be addressed, if you are using these kinds of scenarios for our planning purposes, which which is what is a final goal.

(Refer Slide Time: 40:28)



So, we convert this precipitation time series into what is called as drought indices. So, we use the several drought indices are available, I will not going to details of those. We will use what is called as a Standardized Precipitation Index S P I, which is a indicator of meteorological drought, it is just require the precipitation values, using the precipitation values you convert that into a standardized precipitation index. There are other indices like Palmer drought severity index, Bhalme-Mooley index and then, effective drought index and so on. But, the S P I, is the simplest one that we use for assessing meteorologic drought.

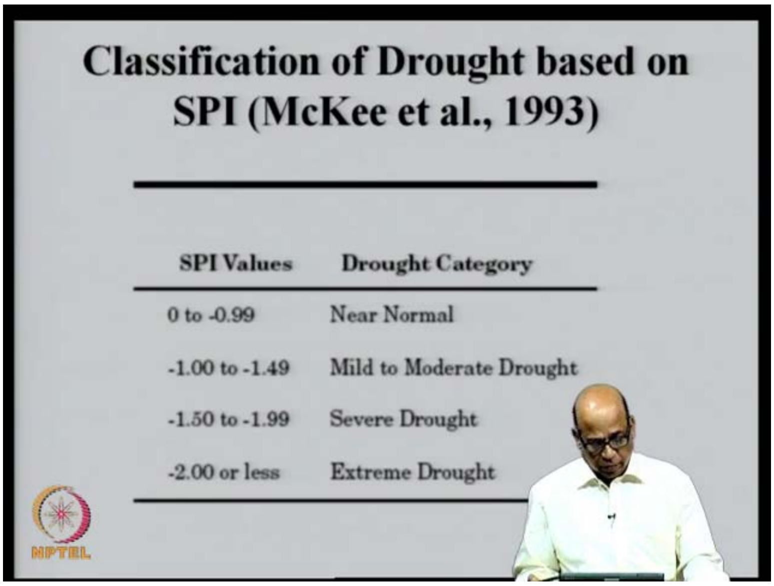
(Refer Slide Time: 41:15)



We convert the precipitation series into the S P I as follows; we fit a probability distribution to the precipitation series. So, this is the annual rainfall here and then, we fit a probability distribution C D F, using our Weibull method or whatever. So, these are Weibull's plotting position and typically annual rainfall as I have mentioned (( )) when I covered the probability distributions, you may fit a gamma distribution. So, the gamma distribution in this particular case fits reasonably good, then we convert this distribution to a standard normal distribution. So, any point here you just transfer it into standard normal distribution and that gives you the standardize precipitation value.

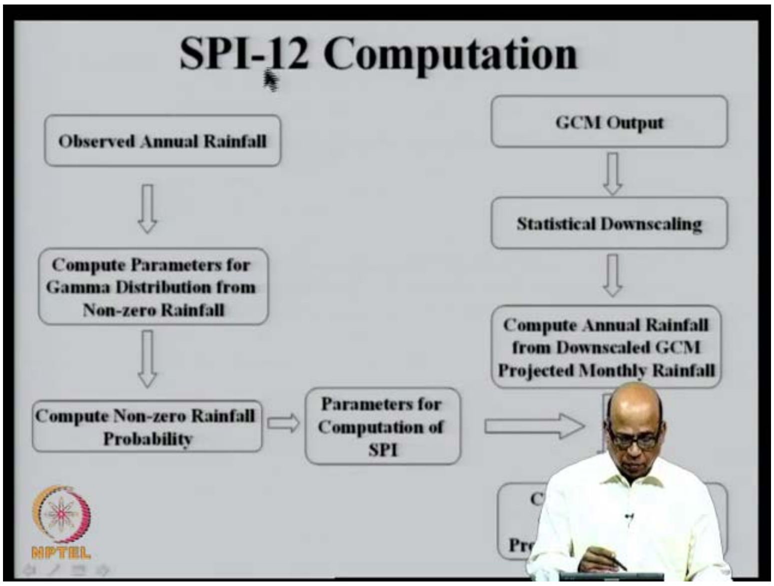
What I mean by that is? Let say that, you start with a particular value here go up, reach the curve and come horizontally reach this point, and this value will be the S P I 12, that is how you convert the precipitation time series into the S P I 12 values associated with various values of this. So, you get a S P I 12 on the x axis here, associated with the standardized normal C D F, starting with the C D F that you have fit for the annual rainfall. So, you get various values of the S P I 12.

(Refer Slide Time: 42:48)



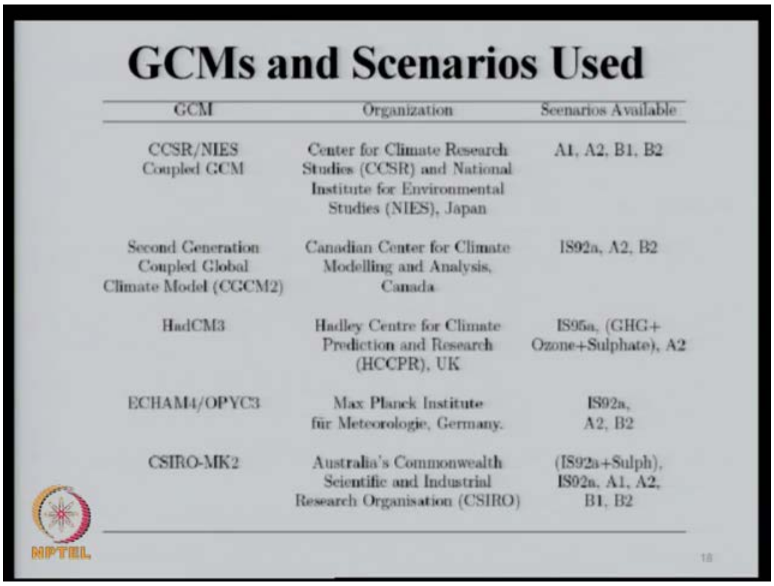
Now, this S P I 12 depending on the value that you get here, this will indicate the drought category. So, 0 to minus 0.99, it indicates a normal situation and so on, mild to moderate drought then, severe drought, and extreme drought etcetera. This is based on the S P I that is standardize precipitation index and the reference is available here Mckee et al 1993.

(Refer Slide Time: 43:18)



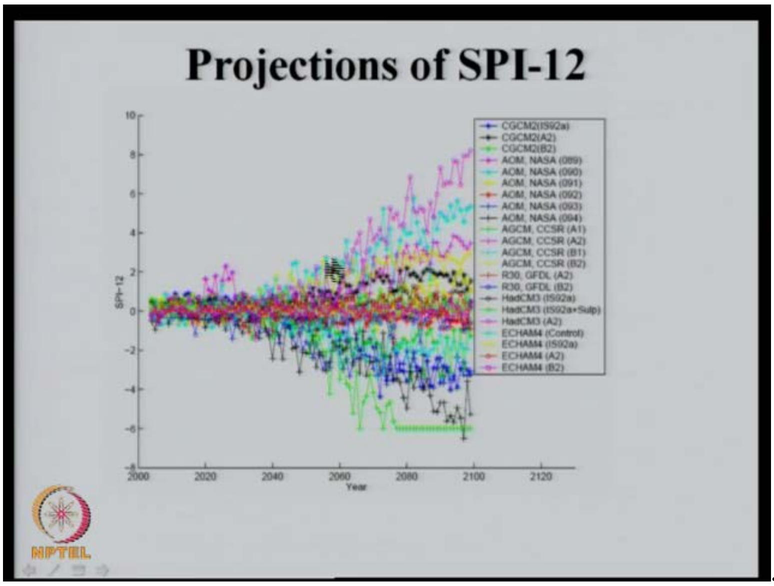
We use I just explain how we obtain the S P I 12 computation, so we do this both for the observed rainfall as well as for the G C M output. So, with the G C M output we have done the statistical downscaling using the regression relationship and then, we have obtained into the future, we have obtained the time series of the precipitation and then, we use we compute the S P I 12 for the future.

(Refer Slide Time: 43:49)



And then, look at how the S P I 12 is likely to evolve into future, when I use several G C M's different G C M's for different scenarios? So, these are the different G C M's that we use for different scenarios. Remember, what I talked about just now was precipitation projection. So, this gives the precipitation projection for a given G C M and for a given scenario, this precipitation projection will then be converted into a S P I projection, Standardized Precipitation Index to indicate how the droughts are drought picture is likely to evolve in future.

(Refer Slide Time: 44:39)



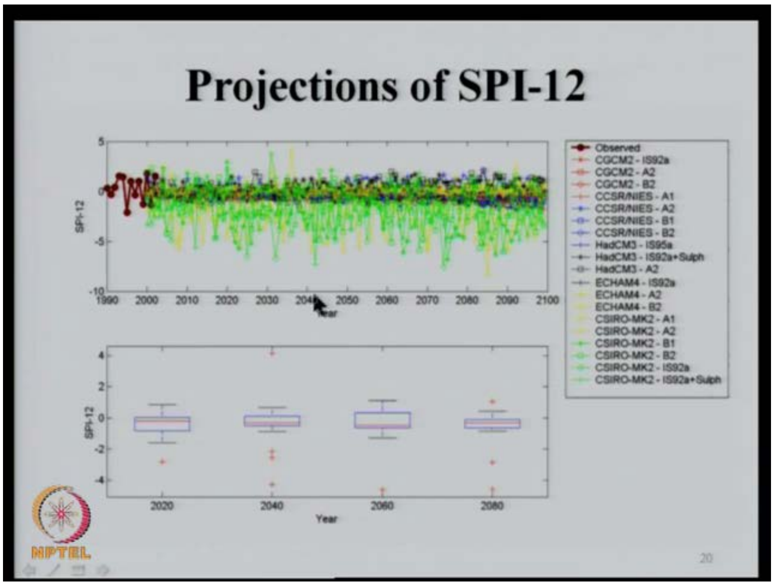
And that we do for several G C M scenario combinations. So, each of them will show 1 G C M scenario combination. For example, this curve may indicate A O M NASA for (( )) scenario, then A G C M C C S R for A 1 scenario, C G C M 2 i s 92 a scenario. So, each curve represents the projection provided by a particular G C M for a given scenario. As you can see here, there is a large amount of uncertainty associated with these projections, even when you are looking at time periods of 2040's.

So, there is a large spread that is being seen here, but this band or this spread starts increasing as you go into the future 2040, to 2060 the band is much higher as you go into 2080, it is much higher and so on. So, the uncertainty propagates into the future and it is important for us to quantify the uncertainty when we want to use these projections for our decisions as well as, develop methodologies and tools to see whether we can reduce the band of uncertainty; because, we know with this kind of uncertainty, it is very difficult for us to use this information into any decision making mechanisms.

One way of doing this is, let me show the same thing we are using the box plots. So, when we have these kinds of projections, we generally provide as I told in the last class, we provide the box plots. So, in 2020 we get a certain type of uncertainty, certain type of range, and certain type of mean median and so on. 2040's we show something, 2060 we show. So, typically the box plots are used to show the uncertainties as you progress in time in the case of projections, using the climate change alright. Now, we look at this, so at various points here 2020's you have a certain ensemble of time series, so this is essentially an ensemble of time series.

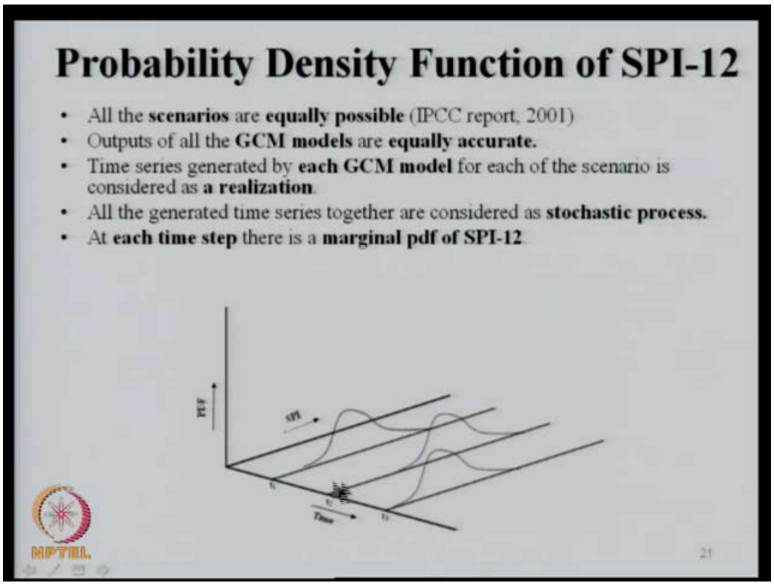
So, at 2020's you have certain spread and certain realizations, and 2040's you have certain different spread here and so on. If you assume that, all this models and all the scenarios are equiprobable; that means, in the future they all have the equal likelihood of providing the right projection; that means, this projection is as likely to occur as this projection, because of our lack of knowledge. We address the first level of uncertainty assuming that, all these projections are equiprobable in future and then, fit probability distributions for various locations.

(Refer Slide Time: 47:44)



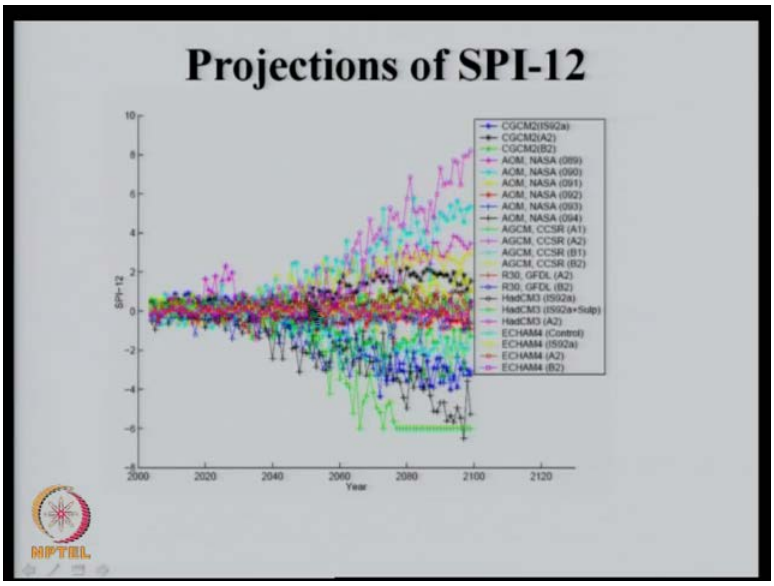
So, you look at 2020's you fit a probability distribution based on the sample that is available at 2020, 2040 you fit another probability distribution based on the sample that is available at 2040 and so on.

(Refer Slide Time: 47:58)



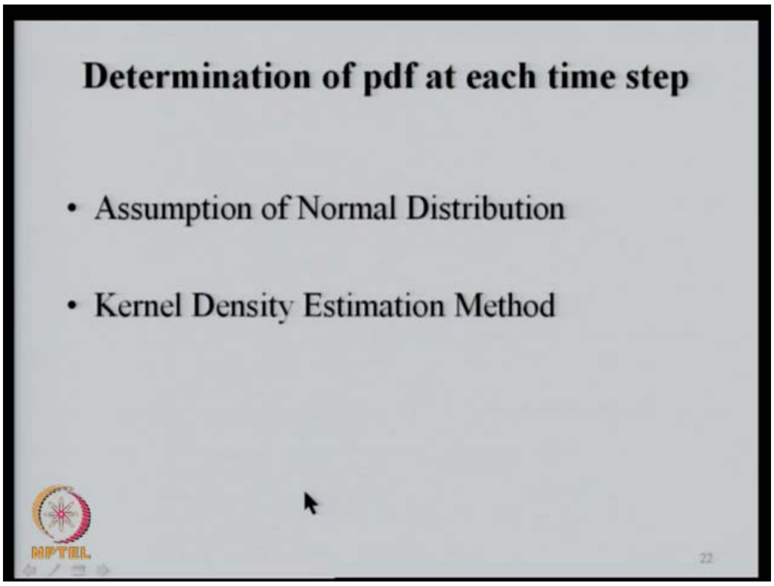
So, you provide the probability distributions of the S P I 12 as the time progresses, this may be 2020, this may be 2040, this may be 2060 and so on. We will then, examine how the probability distributions are likely to change in future, the probability distribution has absorbed or has used all the information that is available through these projections.

(Refer Slide Time: 48:25)



So, at any particular given time, it uses the several time series values that are possible 2040’s to 2060’s let say, this particular time frame, you will use all the time series and then develop the models. So, these are typically done for time windows, not at a particular time 2020, it may represent 2020 to 2040, similarly 2040 to 2060. So, use the ensemble analysis and develop probability distributions at those time windows.

(Refer Slide Time: 49:59)

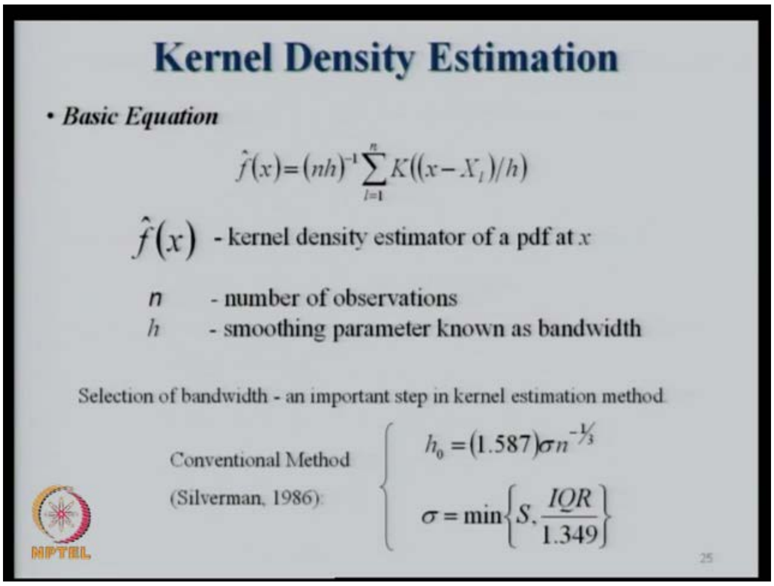


The way we do is, simply assume that every time window it follows a normal distribution and then, get the parameters associated with that, associated with the normal distribution. So, this way you are trying to address the uncertainty, first level of uncertainty you simply say that, all models scenario combinations are equally possible. What I mean by that is the, projection provided by any model scenario combination is as likely to be possible as the projections provided by any other model scenario combination.

And then, you look at different time windows 2020, 2040. 2060 etcetera, you have an ensemble of time series during that particular time period; assume that, this follows normal distribution that means when you fit a distribution to that particular ensemble it follows a normal distribution. And then, you estimate the parameters of that and then provide how the droughts are likely to evolve in future.

The other one will be you use a Kernel density estimation, which I will explain presently and then, come out with the actual probability distribution using the Kernel density function estimates. So, when you use the normal probability distribution (Refer Slide Time: 50:27), you get such picture in this particular case, the near normal condition the probabilities will be like this, the mild draught the probabilities will be like this, extreme draught the probabilities are like this and so on. So, these are obtained from assumption of normal distribution to the various time periods. So, this is 2000 to 2010, 2040 to 2050, 2090 to 2100 and so on. So, like this different time periods you use and estimate the probabilities associated with the mild drought, extreme droughts, severe drought and so on.

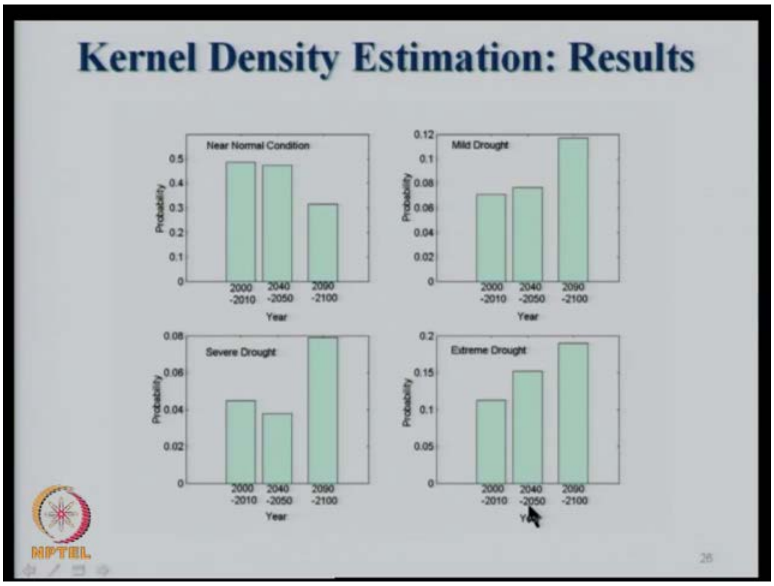
(Refer Slide Time: 51:06)



In the Kernel density estimation, what we do is that, we fit the probability density, this is the Kernel density estimator of probability density, this is the number of observations in our case, it will be the number of simulation that are possible, and this is called as a smoothing parameter and you have the Kernel functions here. Now, this smoothing parameter, which is called as a bandwidth, it is important to assume it is important to assume a certain form of the bandwidth.

And that in this particular case, we assume that the Kernel is Gaussian and therefore, we get this kind of bandwidth here (Refer Slide Time: 51:53), we will not go too much into detail. We just distinguish this between what we get with a normal density as well as, Kernel density estimator.

(Refer Slide Time: 52:02)



The kernel density estimator again can be obtained directly from a matlab subroutine. So, when you do the Kernel density estimator, you get near normal condition like this, and then the severe drought extreme drought probabilities are obtained like this. So, essentially then, what we are doing is, that we are using the downscaling technique to project into the future in this particular case study, I have talked about projection of the precipitation at a meteorological subdivision and specifically the Orissa subdivision.

We project the precipitation into the future, using several G C M scenario combinations; convert the precipitation into a drought index. And in this particular case, we have used S P I 12 the Standardized Precipitation Index, which is computed based on the monthly data and that is why 12 months data comes into picture, because you get a spread, because you get an uncertainty band at various time windows 2020 to 2040, 2040 to 2060 etcetera, your fit probability distributions of the S P I 12 for those windows.

The first level of assumption that you make is, that you simply assume that at each of these time windows, it follows normal distribution and then, obtained the parameters of the normal distribution namely the mean and the standard deviation, for those different time windows. And then, assess what are the probabilities of various levels of drought extreme drought, normal drought normal situation and moderate drought and so on, based on the ranges that are available for the S P I 12.

Next, you use a Kernel density estimator, which can also be done using a matlab program and then, you again assess probability of various levels of droughts. So, this is how you address the uncertainties arising out of the G C M and the scenarios. As I mentioned, our primary hypothesis here is that the projection provided by any G C M is as likely to occur as projection any G C M scenario combination is as likely to happen as projection provided by any other G C M scenario combination; which means, the we were saying that, all this time series are all these projected time series are equally likely to occur in future based on that we address the uncertainty.

There are other other questions to be asked in the uncertainty, before I close I will just explain those things and then we close the lecture. It is possible that, these are all not likely to these are not equally likely then, what we do? We start assigning weights to the projection provided by a certain G C M scenario combination.

How do we assign the weights? We assign the weights based on, how well this G C M scenario combination has performed for the particular region in the immediate past let say, between 1990 and 2005 or 2010, 1990 and 2010 how the particular G C M scenario combination is able to reproduce exactly what has happen to the precipitation in that Orissa subdivision between 1990 and 2010 based on how well it has been able to reproduce we assign weights.

So, different G C M scenario combinations we will have different weights, and we provide what are called as the weighted probability distributions. Then we also address a question of the imprecision in the probability assessment, because of the small sample as I said even if you get all the models and all the scenarios, perhaps you may have about 100 samples, 100 time series. So, at any particular location when you are looking at the probability distribution for a particular month, you have maximum of 100 values.

So, this is a small sample that we are talking about in fact, even 100 you will not have. As I as I showed in the particular case, you may have about 2025 values, and 2025 values you are trying to fit a probability distribution. Therefore, the probability distribution itself will not be precise then, we move onto what are called as a imprecise probabilities. So, we develop imprecise probability distributions and then, provide (( )) of the droughts, that is what is the probability bound for extreme drought, what is the probability bound for normal conditions and so on.

Now, these information, this types of information will be extremely useful in planning for the future. Let say that, you want to plan for agricultural use of water onto the future, so you you must know what levels of probabilities exist for drought in various regions of the country and these kind of information will be useful for that. So, today's was a special lecture on using the stochastic hydrology techniques that we have learnt in the particular course, for the most timely problem of climate change impact assessment.

And as I mentioned in the class in this particular lecture, the climate change impacts the assessment of climate change impacts is burdened with a large amount of uncertainty arising out of the G C M uncertainty, the general circulation model uncertainty, the scenario uncertainty, the downscaling uncertainty and the intra-model uncertainty themselves that is a model uncertainty.

So, it is important for us to use the techniques and the tools that we have learnt in the course, to address all these kinds of uncertainty. And I have provided a simple example of using the G C M outputs to project Orissa rainfall into the future and then, addressing uncertainties by providing probability distributions for different time windows. And these probability distributions will be useful for planning purposes using the probability distributions you can plan for agriculture activities, you can plan for you can convert them into stream flow scenarios and so on, and the plan for reservoir operation. So, developing adopt your responses will require these kind of projections.

So, the next lecture, which is the lecture number 40 will be the last lecture of the course, where I will summarize whatever we have taught we have discussed during all this 39 lectures, thank you for your attention.