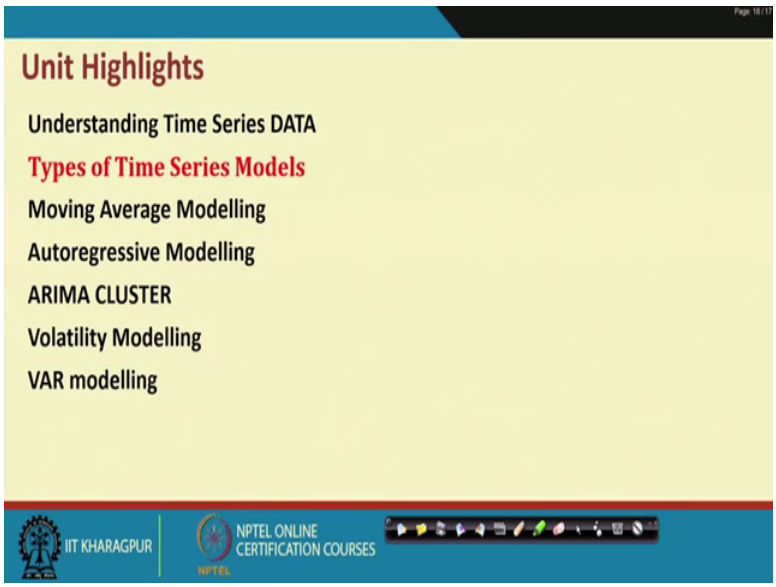
Engineering Econometrics
Prof. Rudra P. Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

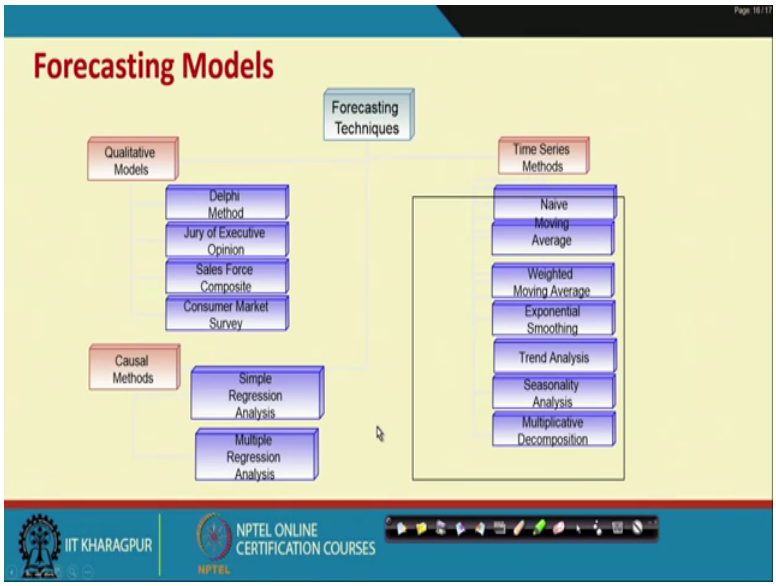
Lecture - 44
Time Series Modelling – Forecasting

(Refer Slide Time: 00:31)



Hello everybody, this is Rudra Pradhan here. Welcome to Engineering Econometrics, today, we will continue with Time Series Modelling in that to we will discuss some kinds of you know time series models, you know with it is specific focus on again forecasting.

(Refer Slide Time: 00:53)



So, let us see what are these you know items which we are discussed earlier? Technically, we have here detailed classifications and that to about the forecasting techniques. So, we have 3 groups or we can say that you know it is a set of 3baskets. The 1st basket is with respect to qualitative models, the 2nd basket is a casual methods, then the 3rd baskets is the time series methods.

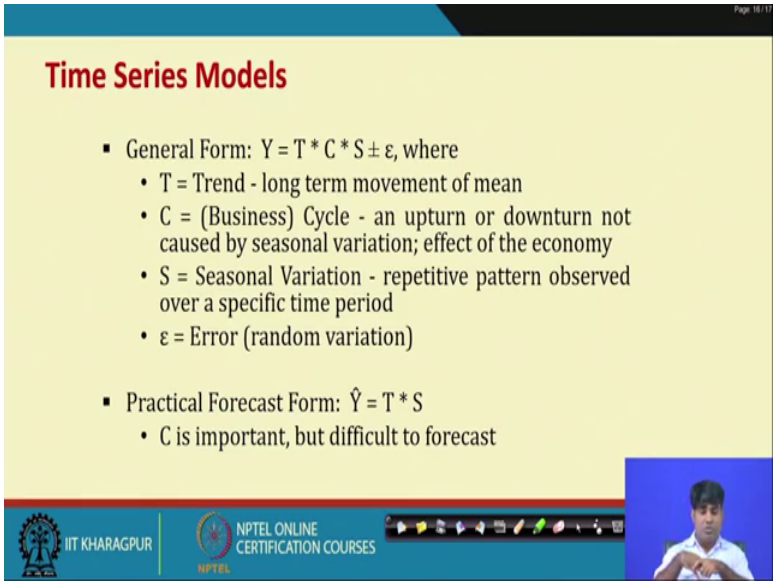
In the qualitative models, we have Delphi technique, then a jury of executive opinion, sales force composite and consumer market survey. And in the case of you know causal methods simple regression analysis and multiple regression analysis. And on the other sides, the time series basket we have naive moving average, weighted moving average, exponential smoothing techniques, trend analysis, seasonality analysis and then multivariate decomposition. In fact some of the techniques which we have already discussed, for instance, causal methods we must discuss that to simple regression structure and then multiple regression structure.

And some of the qualitative you know models, which we have highlighted, of course we have not discuss in details, because for the qualitative models, we need inputs from the a primary survey and on the basis of that opinions, we have to derive some kind of you know information. And that information can be used for you know kind of you know predictions and the kind of no forecasting and as per the particular requirement of you know any you know engineering econometrics.

But, we will mostly focus you know those models, where we have actually data then with the help of you know techniques we can have the estimated model. And then on the basis of estimated model, we can go for some kind of you know forecasting. In fact, in the time series you know models whatever we have discussed earlier, we have gone through some kind of you know moving average techniques.

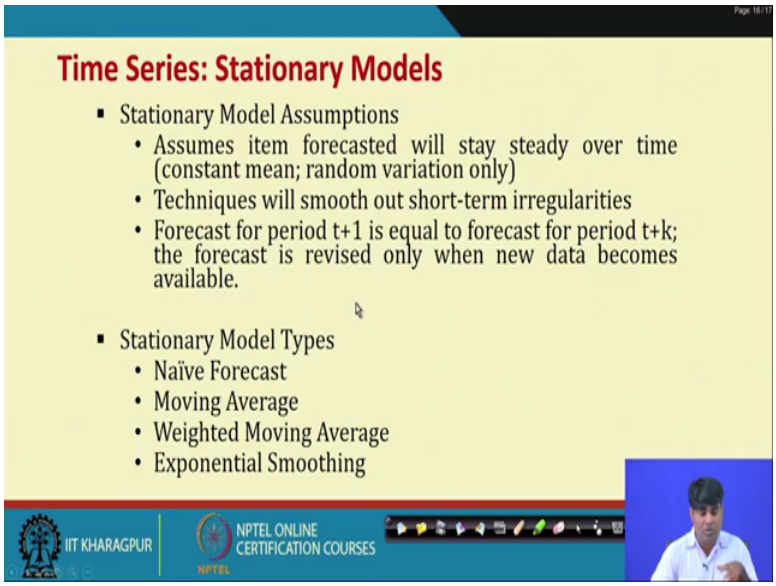
And that too simple you know moving average then you know 3 years moving average, 5 years moving average. And again, we have also discussed the list case mechanisms that is the structure called as you know trend analysis through which you can a you know derive a kind of you know models and that model can be used for the a predictions and the kind of you know decision making process.

(Refer Slide Time: 03:55)



So, let us see you know see how is this particular you know structure. In fact, we have already discussed in the time series models so, we have you know time series components starts with you know T a T trend, then C stands for you know business cycle, S stands for seasonal variations and then there is a error terms. Of course, there is a additive you now structures and then multiplicity structure in any contest. So, these are the various factor through which actually, we like to journalise the you know prediction rules and the kind of you know forecasting rule. And then, we have to go about this particular you know structure.

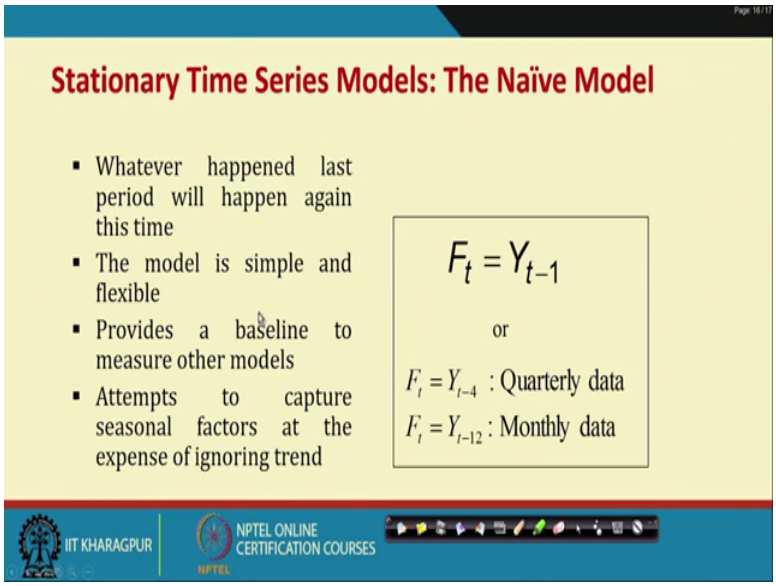
(Refer Slide Time: 04:35)



In the time series stationary models, we have some kind of you know stationary model assumptions. As you know assumes item forecasted will stay steady over time, then constant mean a random variation. Then techniques will smooth out short-term irregularities. And forecast for period t plus 1 is equal to forecast for period of you know t plus k, so that means the kind of you know structures which you like to bring, so that you know the process can be a generalize.

And then we can you know develop a system through which you can do the predictions, and the kind of you know forecasting. So, stationary model types if you look, then we have naive forecast, moving average, weighted moving average and then exponential smoothing.

(Refer Slide Time: 05:37)



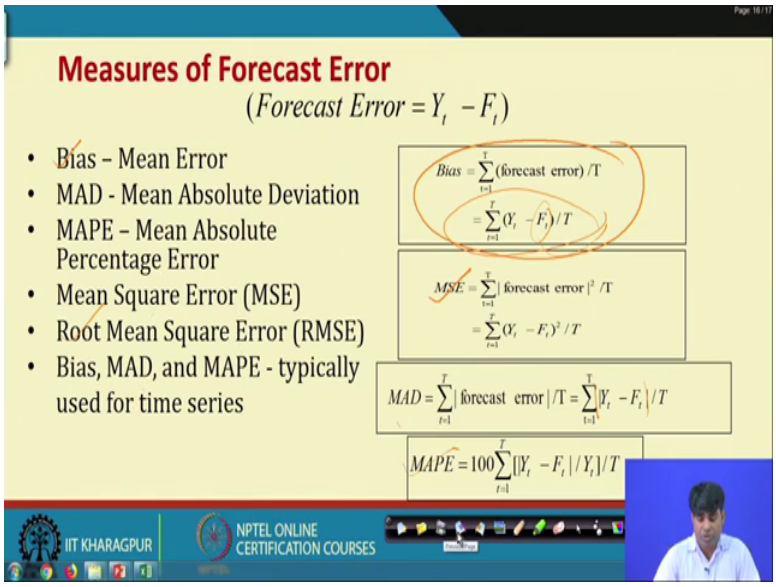
Of course, we have already discussed slightly on this techniques, till we will just highlight once again. And in this lectures, most you know most important is in the focus on the error part that means, technically by using you know a particular methods like you know simple average or you know kind of a moving average, weighted average.

So, we will get estimated you know a kind of you know models that to if say Y is a kind of you know actual variables, then we will have a you know estimated Y, and the difference by default, we will have a error. And again there are various ways you can find you know couple of alternatives and one way to choose the best models to do the forecasting at the kind of you know predictions is the you know what will called as you know the kind of you know error part. And then the error part has it is actually you know kind of you know any kind of you know characters.

Every times when there is a kind of you know comparisons and then the issue is to find out the best out of several alternatives. So, we have to apply this techniques through which actually, we can actually pick up the best one. And whatever you know tools with discuss here right now, like you know min error, min a some errors, so like that you know there are couple of indicators are there. Every time, since it is a error component, every time the choice of a particular model for the prediction and forecasting, exclusively depends upon the items, which are actually you know minimum as per the particular you know requirement.

So, the model is a you know start with like this, F t is the kind of you know quarterly data and then it is you know kind of you know integrated with the previous kind of you know items that is means through historical data. Technically, we call as you know kind of you know lag variables. So, Y t is a current you know variables at a particular point of time by default the t minus 1, t minus 2 is the previous variables that is what you know lag variable 1, lag variable 2 and so on.

(Refer Slide Time: 07:55)



So, we like to integrate all these you know variables while predicting the you know current variables. So, now technically, we have couple of you know indicators here through which we can actually choose the best one out of several alternatives. The first one is the bias that is what is mean error? Which is actually it taken means the difference between true value and the predicted value. Of course, the predicted value is derived from a particular you know process maybe simple average again, maybe moving average in that to 3 years moving or 5 years moving, so whatever for may be.

So, our ideas you know, let us have the predicted you know structure and we have the actual structure. Then the difference will be have the component called as a forecast error. So, now dividing the you know time, then we will have actually mean and a that mean a mean error should we actually you know should be used as a criteria to choose the you know best one right. Similarly, mean absolute deviations we are we like to

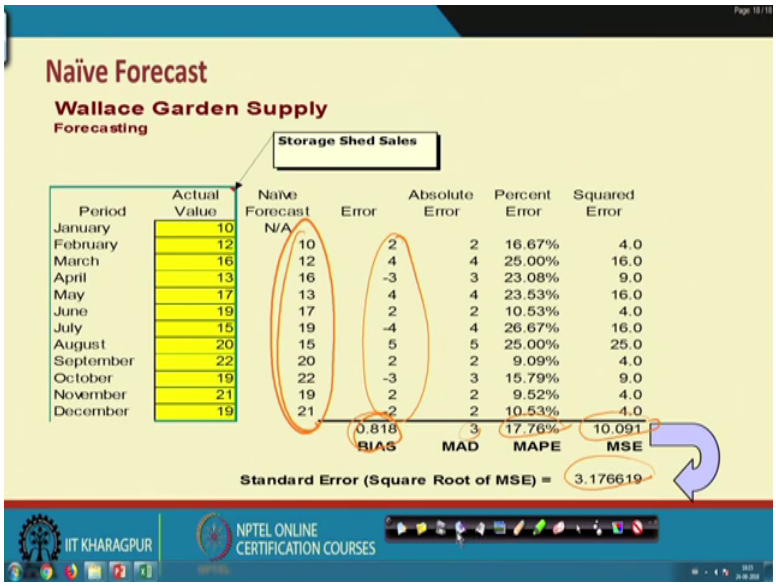
actually a go for you know deviation format so that means, usually when you find out the error term, which is the difference between the actual and the estimated.

Sometimes, the error for you know the error component that to forecasting error, you know positive and sometimes it is negative. So, when you take the modulus, then all the components will be coming as a positive. And again, so that will be the minimum criteria through which you choose the particular you know model. So, it is actually another you know structure called as a mean square error and again we have a root mean square error. So, again forecasting error is having actually a some positive values and some negative values. And when will you sum of, then there is a high chance that you know sum will be equal to 0 that is how the particular line is called as you know bias fit line.

So, now one way to again use this criteria is you to find out the square and that is how it is called as you know mean square errors. So, now if you square all these errors plus and minus, then all will be transferred into a positive quantity, then after that the criteria of you know choosing the best one is the minimum of MSE. Similarly, in the root mean square, so that is the square root of this you know mean square error. So, technically there is you know big difference.

And against there is a mean absolute percentage error. So, it is a again same like you know mean absolute deviation the same way, you have to just divide the actual and then multiplied by 100. So, we will get mean absolute percentage error. And that is the criteria through which you can actually a generalize the kind of you know items through which actually, we can you know predict the a particular you know requirement.

(Refer Slide Time: 11:07)



Now, what we can do, so we like to you know check all these you know kind of you know tools, you know to predict to find out the best ones out of you know variation alternatives. So, let us start with a simple problem, so this is actually monthly data. And let us say you know 1 monthly data of a particular year, so starting with you know January 10, then February 12 like that up to December 19. And we like we like to start with you know naive forecastings and so the actual, and then the naive forecasting will be here you know just you know one point you know a head ok.

So, this is have the kind of you know structure. So, through which actually you can actually find out to the this is what the estimated structures, just you know one point behind so that means, the actuality depends upon the first one. And then the error will be the difference between the actual and the first one so, will find the error component this one. Of course, you know the sum of the error terms should be close to 0 and here a something coming actually 0.82 because of you know some round off issue otherwise, it should be a converse to 0.

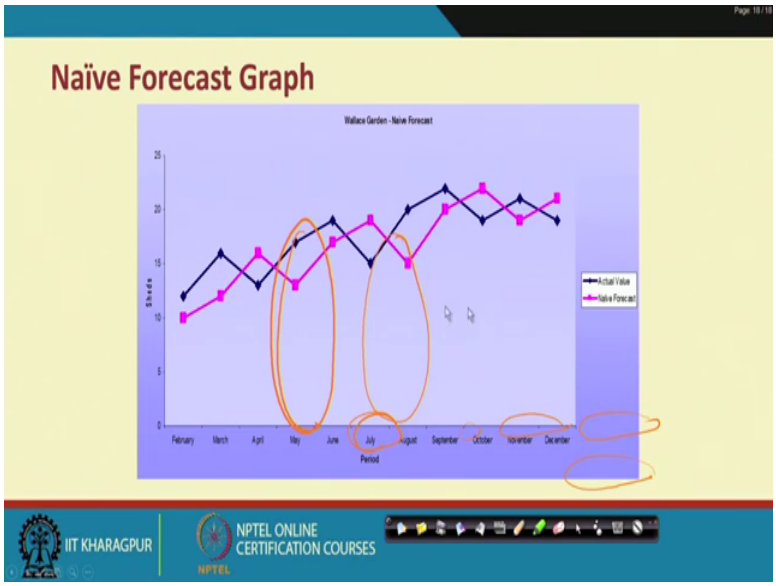
And then if you take the a absolute value, so it is coming 3. And then if you go by percentage error, just multiply divide by Y a actual then multiplied by 100, so you will get actually mean absolute percentage error. Then again, you have the errors you have the error find out the sum, so that is called as a mean square error again. So, take the

square roots, so you will get you know root mean square error, so that means, technically for one model.

So, once you find out the estimated you know model, you can find out the error and then you can check the model accuracy by using the mean error, mean absolute deviation, mean square error, root mean square error and mean absolute percentage error. Of course, they will give some value there is no doubt about it. But, the actual fact is that so, you need actually at least two different alternative models, then the best models which we can choose for from these two, you know on the basis of these you know criteria that is the mean error, mean square error, root mean square error, mean absolute deviation and mean absolute percentage error.

Every time between the two models, where these errors are you know actually low compared to the previous one. And that model can be used as you know indicator for the prediction and forecasting and that is what the benchmark through which you have to you know follow and go head with the you know kind of you know predictions. And so this is what the kind of you know structure and we will go with this particular you know set up to know how we can actually proceed ok so, this is what the case.

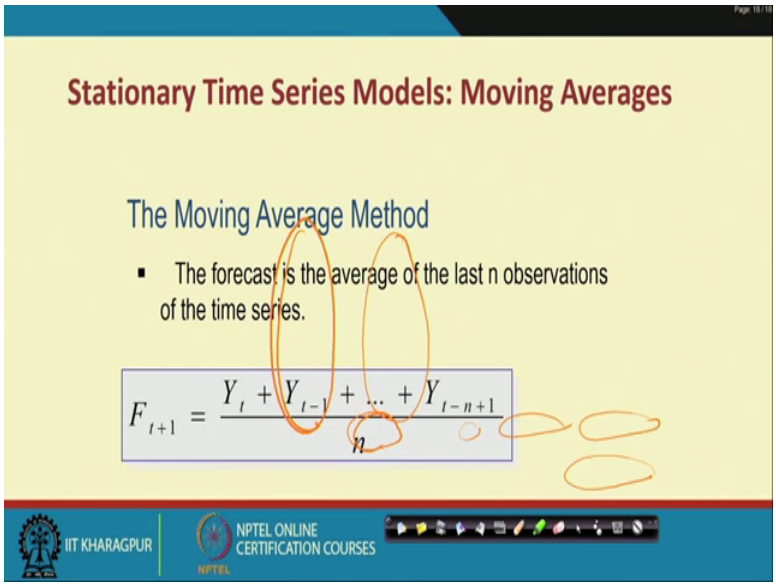
(Refer Slide Time: 14:13)



And we have here actually some kind of you know plotting. So, the a red one is the you know predicted a kind of you know things that is the naive forecasting. And then the black one is the kind of you know a actual one. So, now on the basis of this we have to

find out the difference and you will find sometimes the actual overtakes the predicted and sometimes the expected overtakes the actual. As a result, you will find some positive error and some negative error that is why we have different indicators through which you can actually find out the best ones ok.

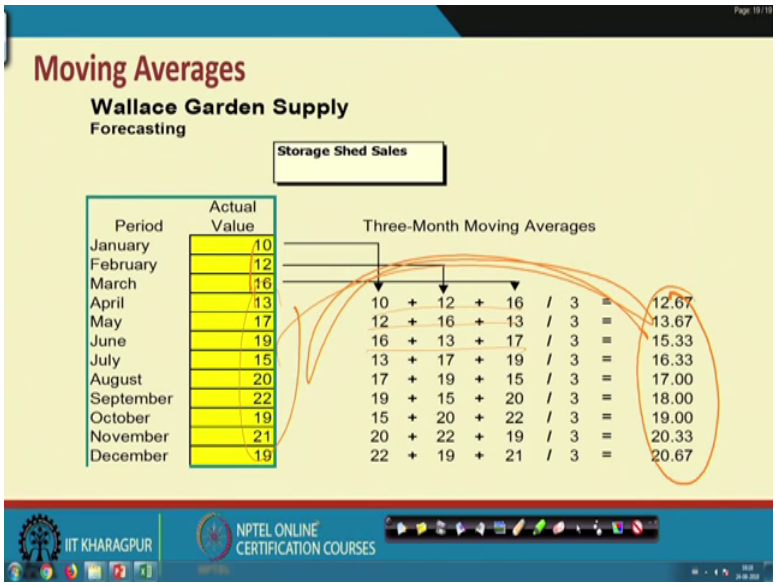
(Refer Slide Time: 14:55)



So, then there is a structure called as you know moving average methods and which we have already discussed. And here you know you have to choose a particular you know moving average mechanism may be 3 years, then the 5 years, 7 years and so on. But, ultimately a with a actual information we must have predicted information, what criteria used that is not the big deal, but the deal is that you know, you must have the actual information and the estimated information.

So, once we can find out the estimated one, then we can find out the error term. And as a result, we are in a position to use these indicators and that to for the requirement of choosing the best one, as per you know any engineering problems or as per the particular you know requirement.

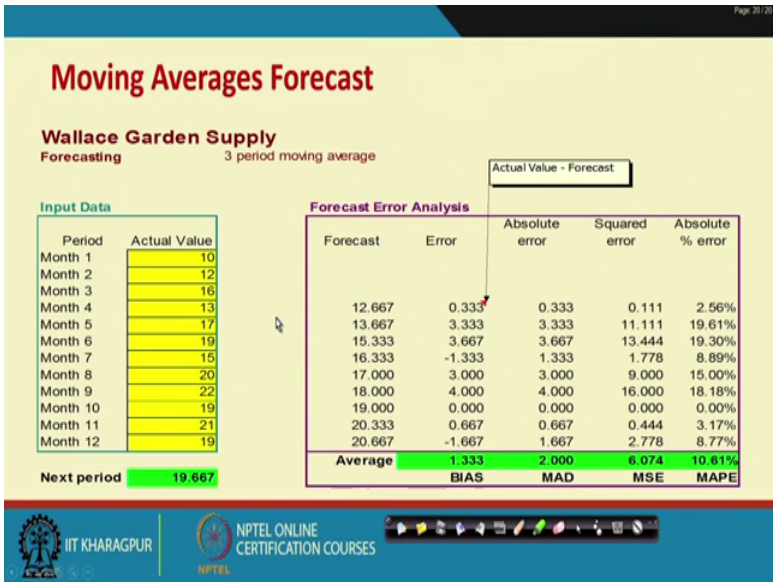
(Refer Slide Time: 15:53)



So, what we can do here, so we can actually check how the moving average all about. And this is again this same data set a monthly data set for a year from January to you know December and here, we are using 3 years you know moving average mechanism. So, just you know first you know what you can do the here, so in fact we have already discussed this one. And so this is the first 3 if it is 3 years smoothing average, then you find out the average again next 3, so you will find out the average.

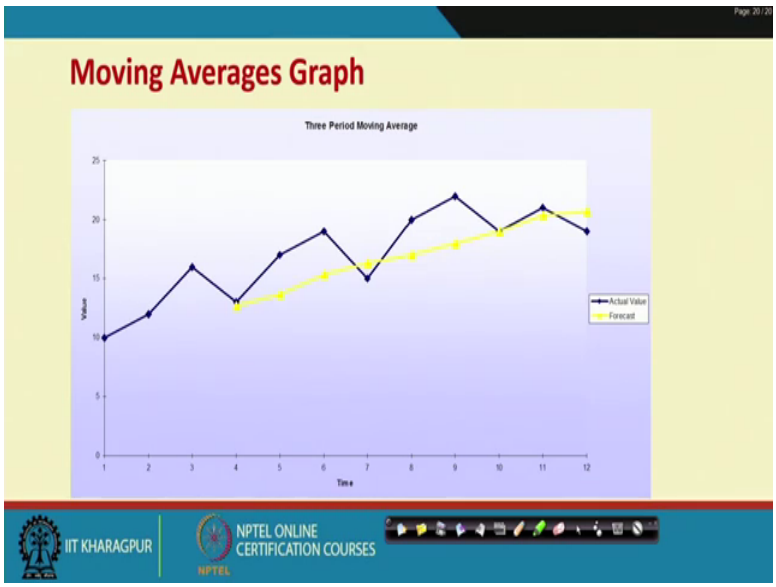
So, like that you know you will find a see here this 10, 12, 16 and then 12, 16, 13 then 16, 13, 17 every time just you know in a kind of you know orders. So, you have to just find out the average of you know 3 items, then as a result this will be your you know forecasted figures. And by default you have the actual figures here and difference between these two will give you the you know error components. Again once you find the error component, you can go you know go for calculating mean error, mean square error, root mean square error and then you find out the best one on the basis of these actually components. So, this is what the moving average mechanism.

(Refer Slide Time: 17:05)



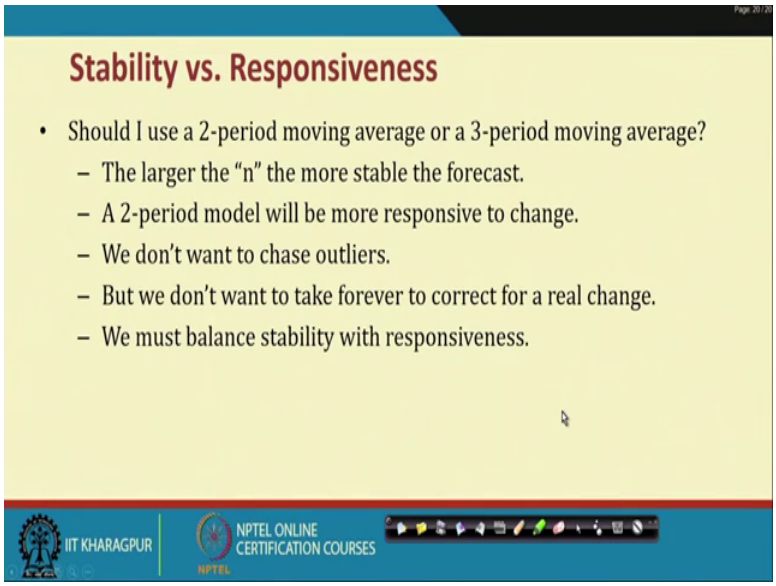
And likewise, we can go for you know yes, you know go for this you know error component, so that means whatever we have discussed earlier. The same things we are here putting the actual value, then the forecasted value, then we will find the error and then you will find different you know indicators by absolute error, by square errors, absolute you know error in percentage. So, these are all again the similar kind of indicators, you can calculate and again so far the best is required, so you must have some alternatives.

(Refer Slide Time: 17:39)



So, again as the graphically the comparison is here, the actual versus the predicted. And again as usual there is a difference between the actual and the predicted. And of course, the predicted will be the good one, if the error component will be converge to 0 in the first instance, so that is what the kind of you know structure.

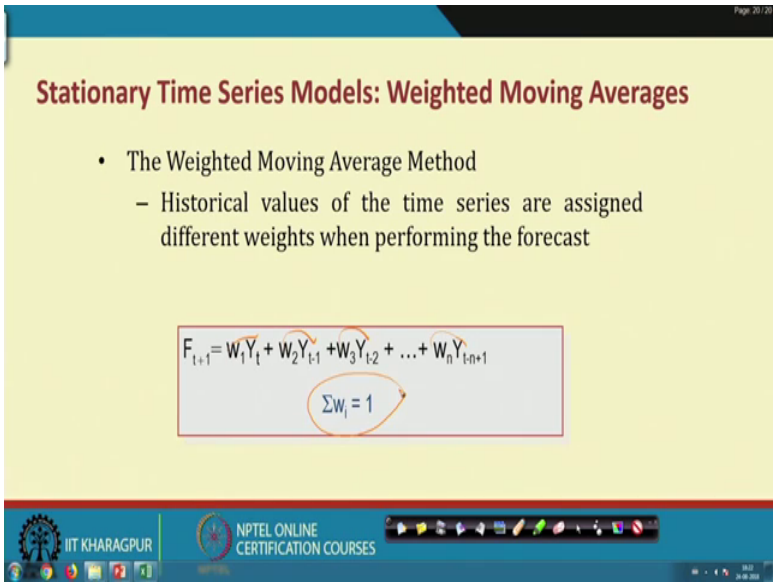
(Refer Slide Time: 18:01)



And again as while doing this kind of you know structuring restructuring, so we will pass through a structure called as a stability versus responsiveness. So, we like to check you know the stability part that means, technically if the model is actually good you know with a particular criteria as a mean error, then obviously to ensure the stability to say that this is actually good ones for the predictions as well as per the requirement.

So, you can use simultaneously the other indicators like you know mean square error, root mean square error, mean absolute percentage, then you finally say that you know yes, that is the model through which you can go for the prediction and the kind of you know management requirement or the kind of you know engineering requirement.

(Refer Slide Time: 18:53)



So, this is another way of you know going through the finding the average you know estimated line, so that is through weighted average. And here actually weight is with respect to different variables, and that to lag of this for existing variable see here. So, it is the game of you know Y t W 1's then W 2 Y t minus 1 W 3 Y t minus 2 and W n T t minus 1 and sum of the W equals to 1, so that means technically the weight W 1, W 2, W 3 will be in between 0 to 1.

So, some points like you know 0.1, 0.3, 0.5 something like that, but the requirement is that you know sum of these weight should be exactly equal to 1. Then, in fact here we have a two different you know package. The first package is that you know you can go for equal weightage principle and the other one you can go for you know unequal weightage principle. Like equal weightage principle means, if there are you know say 3 variables, then by default 1 by 3, 1 by 3, 1 by 3, which does not actually you know give you know very high impart.

But, if you go for you know unequal weight, then it may give better indication and better kind of you know structure to pick up with a things in as per the requirement, but the problem is that how to ensure the unequal weights. If it is equal weight, then depending upon the number of variables just divide the numbers. For instance, if there are 3 variables in the right side divide by 3. So, if it is you say 5 variables case, then 1 divided

by 5. So, 1 by 5 1 by 5 will be used for every weights, so that means, technically W 1, W 2, W 3, W 4, W 5 is equal to simply 1 by 5.

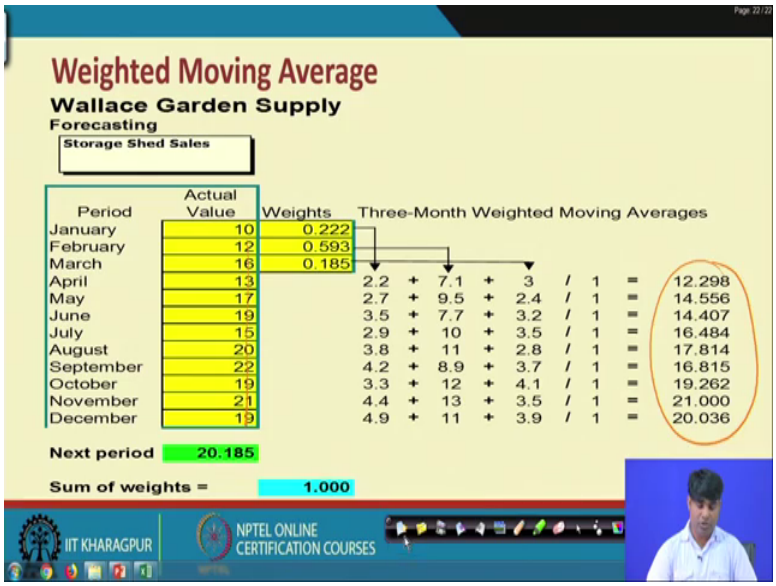
Now, if it is un equal weight, then what you are supposed to do there are various mechanisms through which you can bring that weight. Maybe if the qualitative structure like Delphi technique or some experts opinion through, which we can actually assign the weight vector which may reliable may not reliable. But, mechanically there is a methods through which you can you know justify and that is more practical and more reliable that is what the structure called as you know using factor analysis and specifically using PCA - Principal Component Analysis.

You can you know derive these weights which not which are not necessarily actually as per the equal weight is. So, you will find difference of you know weights and then a corresponding to the particular variable, you can assign weight there. So, it is actually mechanically you will receive these weights through the strata and variables and which may not have any questions or the kind of you know issues.

But, when you will go for equal weights, then there is issue, and means what are the basis you are putting actually equivalent, because the impact of you know lag one variables may not be necessarily equal to the impact of you know lag 2 variable. So, as a result the different lag variables may have a different weights so, assigning different weight is the better choice than the you know uniform weights.

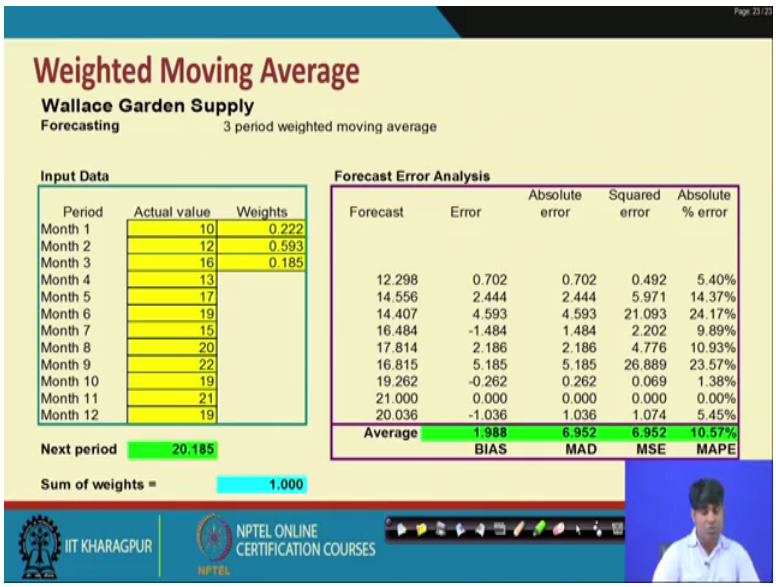
So, in any case we have a two different mechanism through which you can proceed and you know get the estimated model, then do the kind of you know comparison to pick up the best one as per the particular you know engineering requirement. So, this is what the process and then we can actually look for another alternative methods.

(Refer Slide Time: 22:37)



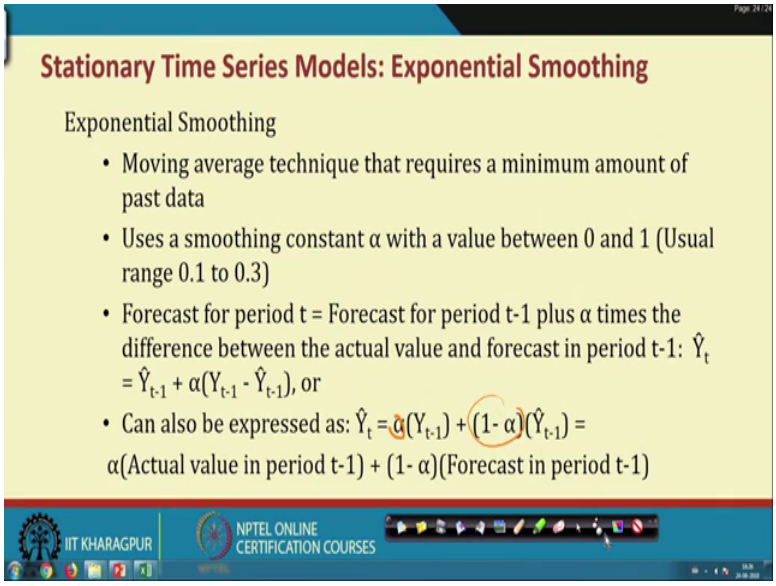
Here you see here the weights and which are actually unequal. So, of course the idea is how to get these weights, which you can actually have from the experts opinions or something like that. You know you can say on the basis of you know principal component analysis so, that means technically it is not the big deal again. So, base on 3 as moving average so, now instead of you know just adding 10, 12, 16. So, you have to just multiply 10 into 0.22 and then 12 into 0.59 plus 16 into 0.19 then you get the average.

(Refer Slide Time: 23:35)



Similarly, you have to follow every steps, then you will have ultimate land of the predicted values and where we have the actual value the difference by default will give you the error component. So, that means it is a more or less actually same structure just you have to follow you know little bit you know kind of you know different you know setups through, which actually you can get these results. And that maybe the good one compared to the previous one that is why we need actually the kind of nervousness, structure and stability structure where different methods we are applying. And then checking that you know the results are uniquely actually a you know deciding factor through which you can do the forecasting as per the particular you know requirement.

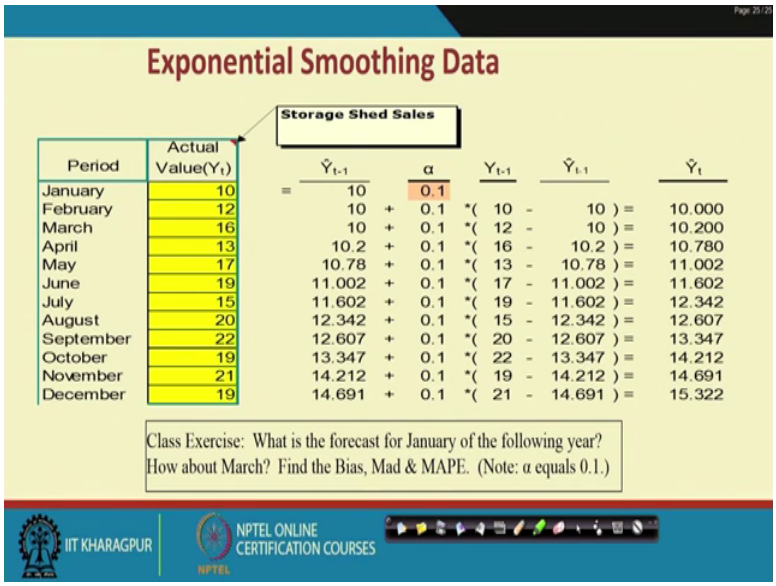
(Refer Slide Time: 24:05)



Then there is a time series models exponential smoothing. And here the same structure we use lag variables and again, so this is same weight you know like we have discussed here. Then here alpha is a weight vector, then by default 1 minus alpha is the another weight vector. Since, the some of the weights are actually equal to 1 so, if you find weight of a particular vector, through by some clue then by default.

If it is a two variable case, the other one will be 1 minus alpha, but if it is actually more than two variables, then these criteria may not be very you know easy to apply. But, in that case you know if you mode number of variables, so the best idea is actually to use pc, and have the weights, and then do the kind of you know estimations and the kind of you know forecasting as per the particular you know requirement ok.

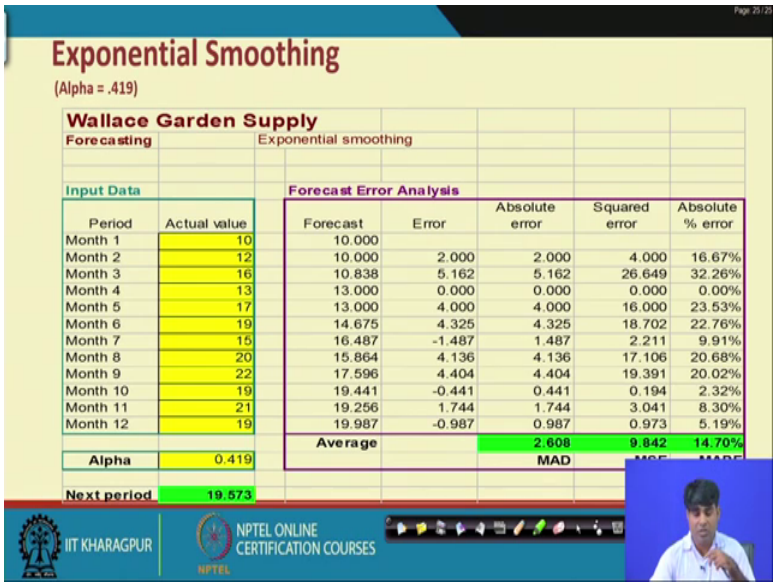
(Refer Slide Time: 25:03)



So, this is what the process and then you can find out the see here the exponential smoothing. So, you have the actual variables here and then this predicted variables Y t minus 1 and then alpha of Y t minus 1, so and so that you know the weight can be you know used this is for the alpha Y t minus 1 and 1 minus alpha Y t minus 1. So, this is how you have to actually just you know that means this plus this is weighted component and this single one.

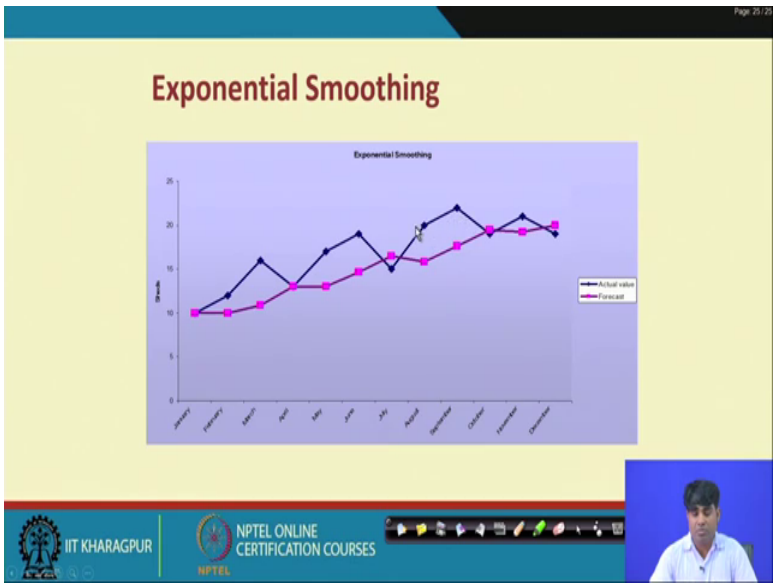
Then finally, we have the estimated coefficient and the kind of you know estimated equations. And then we have a actual Y and the estimated Y again we will find the error and check whether the error is actually maintain this (Refer Time: 25:46) by deploying different methods like you know mean error, mean square error, root mean square error and so on.

(Refer Slide Time: 25:55)



So, these are various techniques through which actually you do the kind of you know structuring, restructuring, get the estimated value, do the predictions, check the reliability, check the validity, check the stability. Then finally pick up the good one, which can be really useful for the particular you know engineering requirement.

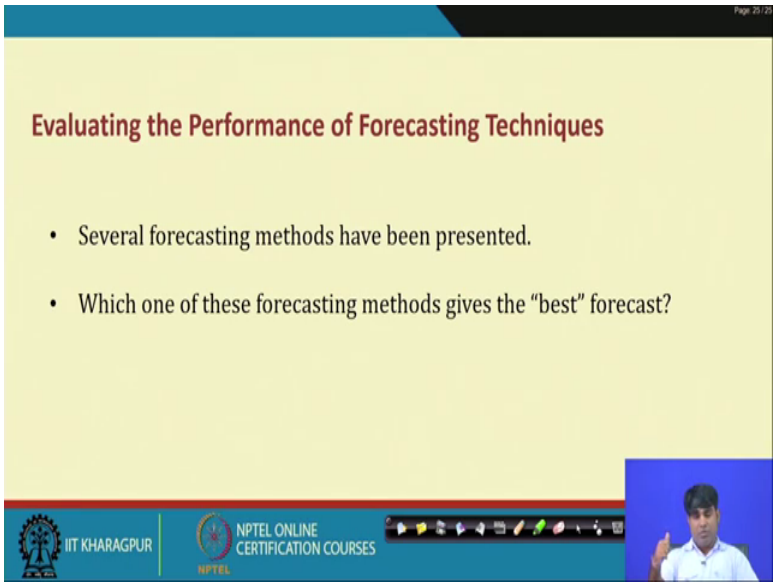
(Refer Slide Time: 26:15)



Again so, every times you can at in a tabular format, you can check the kind of you know stability, and the kind of you know error component. And then again you can you know plot and then you can you know check the kind of you know best ones as per the a

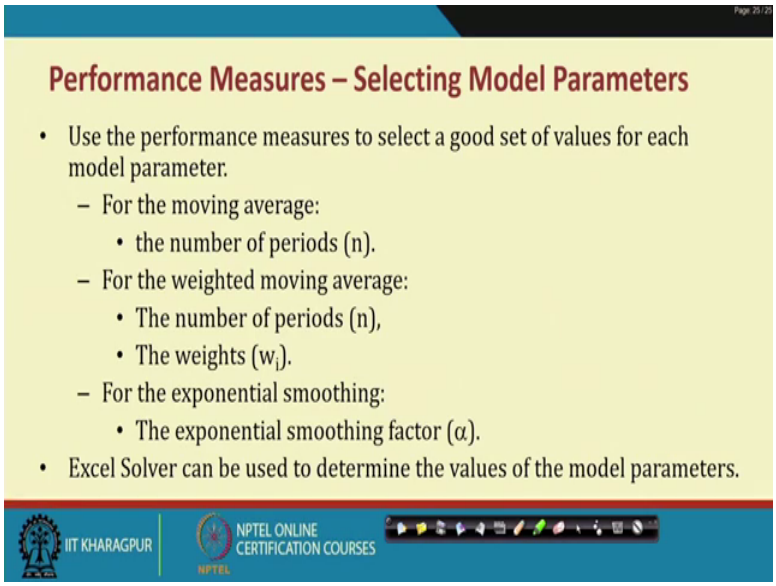
particular you know requirements. So, this the graphical plotting here the red mark line is the forecasted line and the black one mark black one is the kind of you know actual line. So, the difference by default will have actually the error component.

(Refer Slide Time: 26:53)



So, now so the evaluating you know the performance or the forecasting techniques. So, we have you know several forecasting methods which we actually frequently use to do the particular you know forecasting. Of course whatever methods we have discussed in these lectures are specifically relating to a one variable case, where you have a one variable. And by default we are creating a additional variable with the actual data or the original data by introducing the lag variable. So, you can create n number of lag variables while validating the particular you know process.

(Refer Slide Time: 27:33)



And then we have to check we have to choose which one is the best as for the particular you know requirement. So, likewise, there are some other methods through which you can actually do the analysis and then have the estimated model, then check whether the estimated is perfectly fit by deriving the error component, then use time series indicators like mean error mean, square error, root mean square error, mean absolute deviation, mean absolute percentage error. So, that means these are the typical time series indicators through which we can you know check the reliability of the models, stability of the model stability of the model.

And then we come to the conclusion that you know this is the model through which you can you know do the prediction, do the forecastings, and as per the you know as per the particular you know engineering requirement and that too in a kind of you know decision making process. There are certain other you know models are there which we can also add value to this particular you know selection process, which we will discuss in the next lectures. With this we will stop here.

Thank you very much.