Remote Sensing of Leaf Area Index and Primary Productivity

Prof. M. D. Behera
Centre for Oceans, Rivers, Atmosphere and Land Sciences (CORAL)
Indian Institute of Technology, Kharagpur

Lecture - 15
RS of Primary Productivity Methods - LST, ML, Inversion/ Passive RS

Welcome back. So, we have been discussing about the different methods with respect to primary productivity assessment using satellite remote sensing. So, we discussed with respect to the empirical methods particularly the vegetation indices based and then, we went to the process-based models, and we also discussed the SIF based method in the last class, the lecture number 14 in detail.

So, we will touch upon the other three in terms of the land surface temperature, the machine learning approaches which are very important these days and lot of works are going on in terms of remote sensing based primary production estimates and also the inversion.

So, inversion both in terms of vegetation and atmospheric inversion gives us a very good estimate with respect to the global scale carbon flux. So, understanding this will perhaps give us a very detailed understanding with respect to the different methods available in terms of remote sensing based primary products and estimates.

So, let me count it again. One is the empirical based methods with respect to different vegetation indices and many other variables. Second process based: where it could be bio geophysical, biogeochemical process-based models so, that is very important.

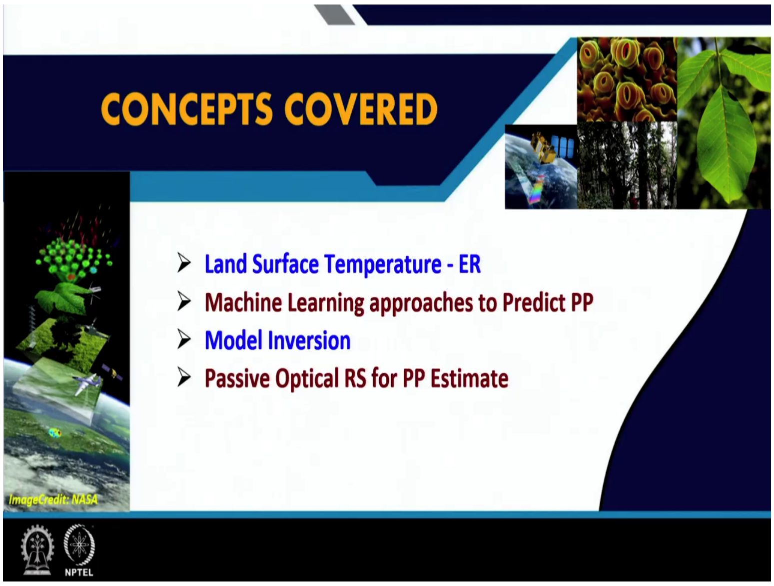
Third is the SIF what we understood in the last class lecture 14 is a more realistic in contrast to the others which measure a kind of potential. So, it is a more realistic in terms of solar induced fluorescence-based measurement. And we will see the leave the land surface temperature which gives vital input as far as a proxy of respiration is concerned. So, that is a very important parameter in that way and understanding it also will help us in terms of giving detail inputs in our climate, or land system models or DVM's.

And also, and these days in the computing world, we are flooded with lot of models in terms of machine learning, artificial intelligence, neural networks so, where we integrate many more proxies of the vegetation variables and non-vegetation variable which have a linkage or in terms of contribution in product means growth and production.

So, we also integrate that using the ML approaches machine learning approaches and inversion as I mentioned at a very big scale or a global scale that gives us indication with respect to the fluxes of carbon dioxide. Also, the on other one which I am not included in our discussion is the eddy covariance-based flux measurement, though it will come, but I purposefully did not include because of paucity of time, but you can still understand it, the literature is available anywhere and everywhere.

So, and with respect to the time, we will try to conclude or summarize some of the things with respect to the passive remote sensing based primary production estimation.

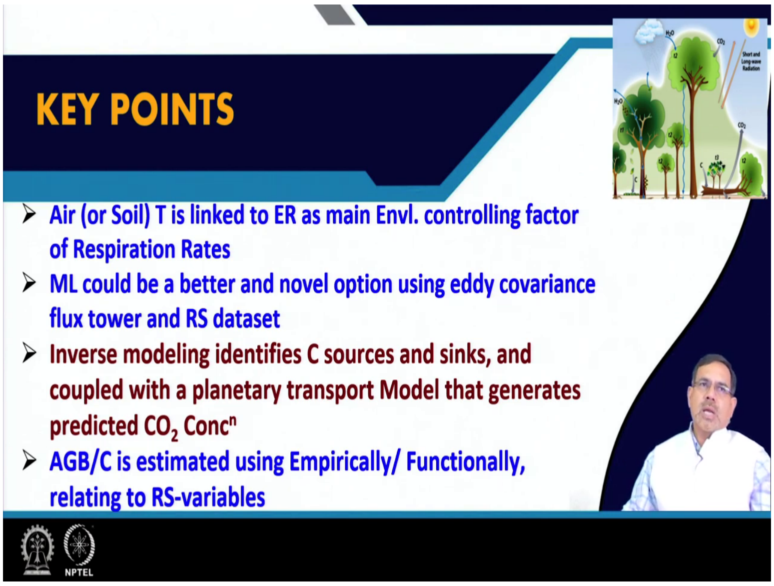
(Refer Slide Time: 03:47)



So, let us move on with respect to the concepts covered. So, the land surface temperature that gives clue with respect to the ecosystem, respiration and machine learning approaches to predict assess estimate the primary production and model inversion where lot of works are going on and still in a preliminary stage, but it holds good and needs to move a long way as far as the global scale carbon flux is concerned.

And we will summarize with respect to the passive optical remote sensing based methods for primary production estimates.

(Refer Slide Time: 04:32)



So, coming to the key points, we will cover are the temperature which both soil and air is linked or well linked to the ecosystem respiration because the temperature or the land surface temperature is a main environmental controlling factor of respiration rates.

Second the machine learning approaches could be a better and novel option using lot of data including remote sensing data sets and also could integrate the data from different sources including eddy covariance-based flux measurements or tower-based flux measurements.

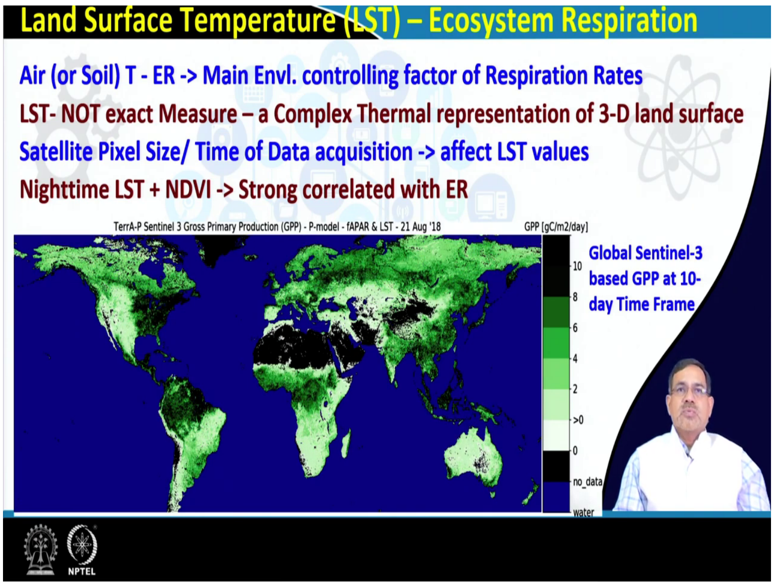
Then third, inverse modeling which identifies the carbon sources and sinks. So, in atmosphere, the carbon means say it unlike in some other place it gets degraded in terms and which we can measure in terms of C 12 and other isotopes. So, here, this in atmosphere, this is everything is coming or getting released from different sources.

When we understand the carbon story, we will know that they are coming from different sources as far as from the carbonate rocks, the marine, the terrestrial, the soil and many other sources and all other sources. But here in atmosphere, the only small or minute source is some amount of carbon addition with respect to the movement of the airplanes.

So, inverse modeling that is why identifies the carbon sources and sinks and can couple with the planetary transport model so that is very important and complex and also complicated. So, coupled with planetary transport model that generates and predicts the CO 2 concentration. So, the latest IPCC reports are including more and more case studies from on the reinforced modeling as far as the global carbon flux is concerned.

The fourth AGB or C so, B stands for Biomass and C stands for Carbon as we know 50 percent as a thumb rule of the biomass is carbon. So, above ground biomass or above ground carbon is estimated very traditionally using empirical or functional relationship and mostly from the optical remote sensing variables that we will try to summarize.

(Refer Slide Time: 07:22)



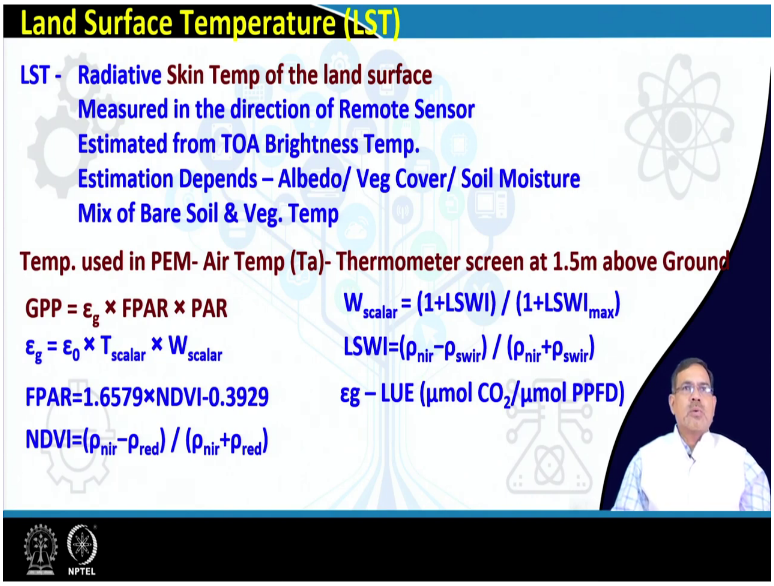
Here is a good image and in motion, we can see the image from the global Sentinel 3 based GPP at 10-day time frame. So, and the scale is there with respect to GPP in terms of grams of carbon per meter square per day and we can very well see it for the different time period it is similar; it is displaying.

So, the LST is does not measure exactly the air or soil temperature, it is actually a complex thermal representation of a 3-D land surface as far as the remote sensing-based measurement is concerned and the satellite pixel size and time of data accusation is very key, and it affects the LST values.

Today and this time so, let us say 16th July and rainy season here over IIT Kharagpur and this time let us say about 12, 12:30 and if the satellite takes the measurement in terms of the LST, this will not be the same in terms of the December 12:30 PM. So, the time or time and the place and so, that is very important so, that has to be kept in mind.

So, again the nighttime and the day-time measurement with respect to LST involves lot of complexity, but researchers have found good relationship with respect to the combined nighttime LST plus NDVI with the ecosystem respiration.

(Refer Slide Time: 09:19)



So, let us move on with respect to our understanding. The LST as we know it measures the radiative skin temperature of the land surface and in the direction of the remote sensor. So, the radiation in the direction of the remote the sensor. So, it estimates; it is estimated from the top of atmosphere brightness temperature.

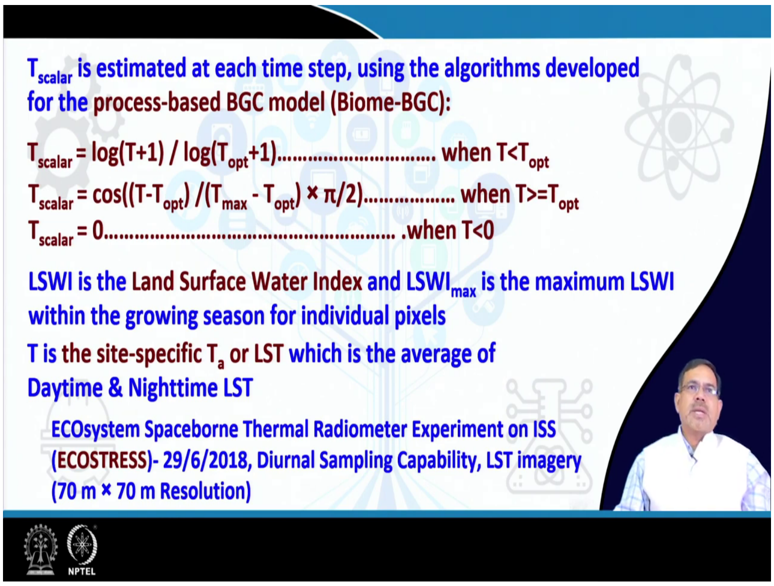
And this estimation depends on lot of things including the Albedo that means, the vegetation, the land cover from where you are taking the measurement, what is that land cover it is a bright soil or it is a water body or it is a muddy area or it is a vegetation forest or grassland so, that is very important and what is the moisture level. So, many things actually complicates what you say estimation in terms of the radiation of the skin temperature.

And more importantly the mixed pixels or the mixture of bare soil, vegetation and other things so, mixed pixel is very important. So, that is also dealt with the on-mixing algorithm in a different way. So, our understanding is very clear that LST is measured as radiative skin temperature of the land surface and there are complexities.

So, the temperature used in the production efficiency models and the air temperature so, that is actually the thermometer screen at 1.5 meter above ground and we can estimate it in terms of GPP, I link it with respect to GPP where you have the g that is the light use efficiency factor, FPAR and FR, PAR and FPAR.

So, this g has the scalars that is the temperature scalar within it. So, T scalar indicates or involves the temperature scalar component which is taken or can be included or integrated as far from the remote sensing or satellite derived variable as in the spatial scale across the globe.

(Refer Slide Time: 11:40)



So, the T scalar is estimated at each time step using algorithms developed for the process based BGC or biome-BGC bio geochemical models and can be represented with respect to these equations at different steps of the estimation. So, T is the site-specific or the LST which is the average of daytime and nighttime LST.

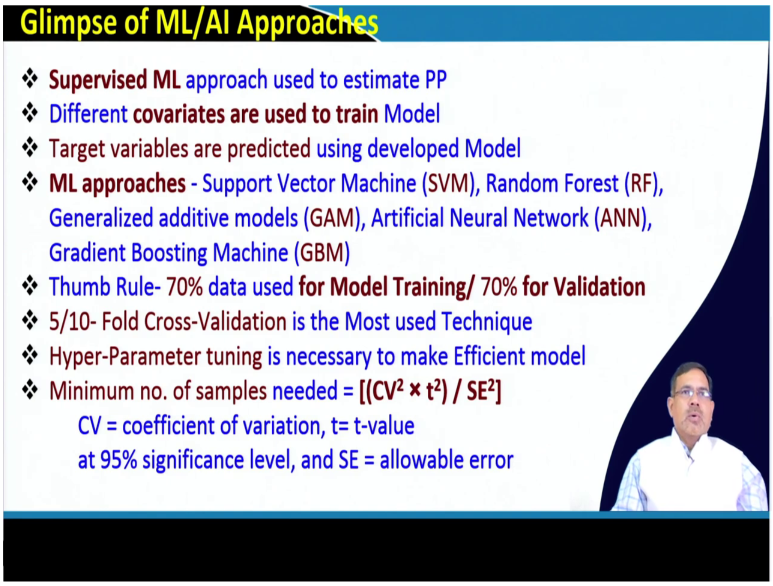
So, friends this understanding of LST is very important and these days lot of models are integrating the satellite derived LST to fine tune or finally, assess the respiration or you say the ecosystem respiration component. So, we must understand the clue why we need or where the LST contributes in terms of GPP estimates or assessment.

So, a recent development since 29th June 2018 more than 3 years back, the eco-stress sensor you might be you might have heard which is mounted on the ISS International Space Station platform which is expanded in terms of Ecosystem Spaceborne Thermal Radiometer Experiment.

So, it has a good capability in terms of diurnal sampling day and night so, you take care of the component in terms of the LST both and day and nighttime LST and it is giving us at a very high resolution in terms of 70 meter by 70 meter.

So, we must use the ECOSTRESS data for assessing this. If some of you are involved in this may try with the ECOSTRESS data, but we have to be very careful that this ISS platform it is at around 300 kilometer from earth and we have to look at the data whether it is available in a regular or you say the data complete to a in terms of the earth surface is concerned.

(Refer Slide Time: 13:59)



So, let us move on to quickly to the next approach what all of us are, many of us are doing these days in terms of the computational world. So, glimpse of the machine learning or artificial intelligence approaches. As all of us know the so, supervised machine learning approach used to estimate primary production and there are different covariates used to train different machine learning models and based on that, the target variables are predicted using the models which are developed.

So, in example we know few names in terms of their common use so, very commonly we use the random forest what you say model in term and SVM's Support Vector Machine, GAM which is an additive model in terms of generalized IT model, ANN, GBM so, many more in the lineup. So, there are lot of machine learning approaches as we have been discussing, they give us the flexibility to integrate many more correlates or variables that could be your proxy in terms of the primary production estimates.

And in these models mostly in terms of data integration or data used as a thumb rule, 70 percent of the data are used for or are used to train the model and rest 30 percent I am sorry it has to be 30 percent so, 30 percent of the same data has to be kept in terms of validation as a principle.

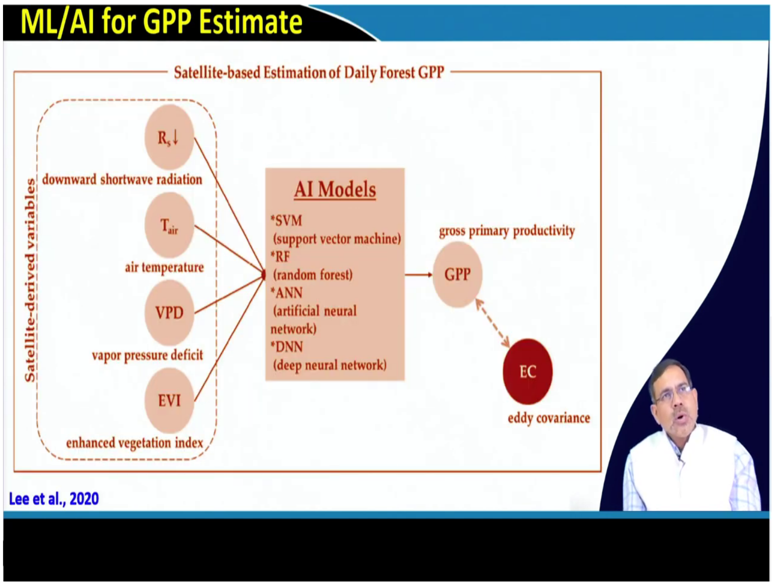
So, whatever data you have based on a data splitting methodology, we have different program, software's and algorithm for that. You separate the data between 70 and 30 ratio and 70 percent data one can use for model training whereas, 30 percent can be used for model validation or validation of the output.

And about 5 to 10 percent fold cross validation is mostly used for this and hyper-parameter tuning is necessary to make the model more and more efficient. So, the minimum number of sampling or the samples are needed for the model to good get a good results and these could be a dependent with respect to the coefficient of variation t value and the significance level and what kind of error or the error which is allowed to us.

So, these are we need to understand lot of with respect to the statistics, with respect to data what you are using and the result what we are getting. So, this makes our job little comprehensive in terms of getting a good result because if we integrate lot many variables and which are not related or done contribute just will give a statistical good or bad or some estimate, but it could be meaningless.

So, we have to be very careful that what are the things or what are the variables that we are considering for the machine learning based model.

(Refer Slide Time: 17:10)

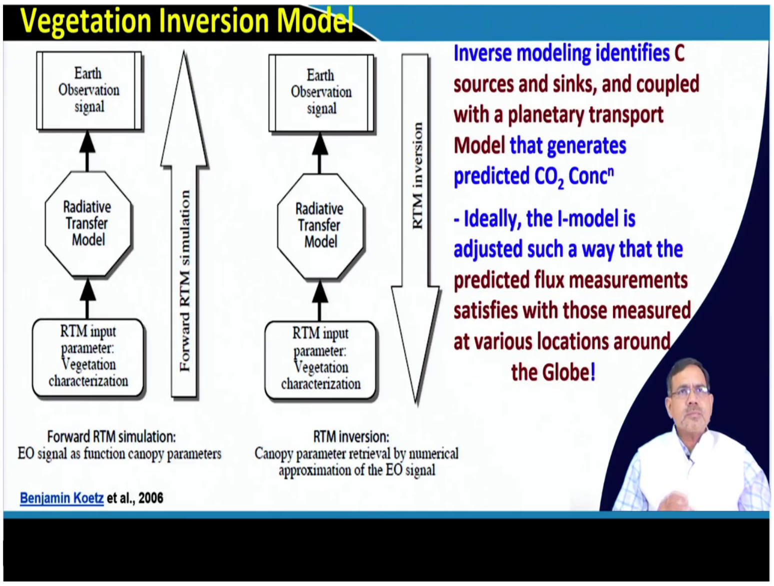


So, a schematic diagram taken from Lee et al., 2020, satellite-based estimation of daily forest GPP or daily forest primary production. So, satellite derived you can see the downward shortwave radiation RS, air temperature, vapor pressure deficit, enhanced vegetation index and many more you can integrate using AI model.

And finally, you get an output with respect to a GPP and that you can validate with another set of data what we have been discussing with respect to the eddy covariance based flux measurement or the primary production estimate for any ecosystem. So, this kind of studies have become a rule these days and more and more studies are coming up to understand the carbon or the primary production of at ecosystem level.

At ecosystem level by integrating more and more variables and by fine tuning their relationship and the contribution optimal contribution and as you know the flux tower data helps in validation because this is a totally independent data source or source of data as far as the production is concerned or the primary production is concerned.

(Refer Slide Time: 18:39)



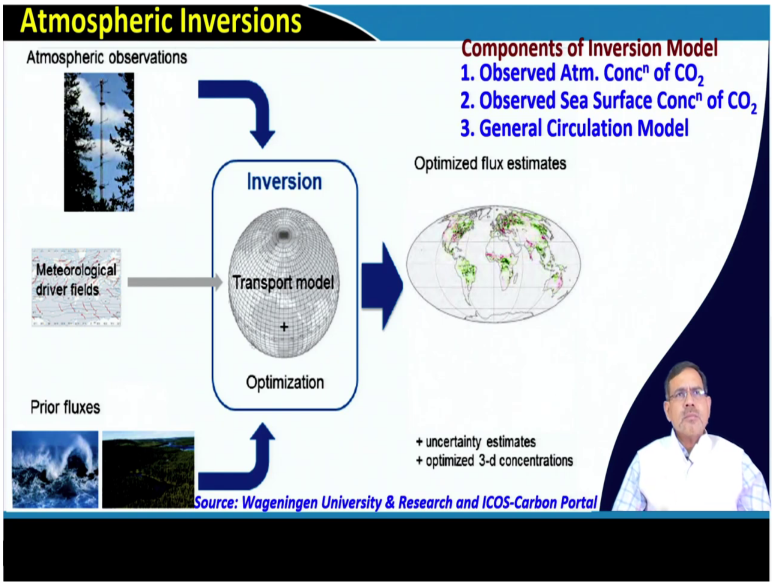
So, let us touch upon the next one as we have been discussing with respect to the inversion. So, we know the carbon dioxide flux or more precisely the carbon flux in the atmosphere so that we can model in a forward RTM simulation or a downward RTM inversion.

So, in the schematic diagram on the left-hand side, we see the RTM input parameter and vegetation characterization goes to RTM and then, observation signal it is considered as a forward RTM or radiative transfer model simulation and the earth observation signal as a function of the canopy parameters which can be inverted or which can be inverted as has been shown on the right-hand side diagram.

So, this inverse modeling identifies carbon sources and sinks as we discussed and coupled with a planetary transport model that generates predicted carbon dioxide concentration. Ideally, what happens? These inverse inversion models, they adjust in such a way that the predicted flux measurements, carbon flux here satisfies with those measured at very at various locations around the globe.

So, these various location measurement is very important which mostly comes from the satellite based measurements or from the flux tower based measurements from different locations including over the vegetation space, over the marine space, over many other land cover space which have a bearing with respect to carbon dioxide flux. So, this inversion as we discussed is means as we are understanding in terms of a broad sense, it gives us a very good estimate in terms of global carbon flux.

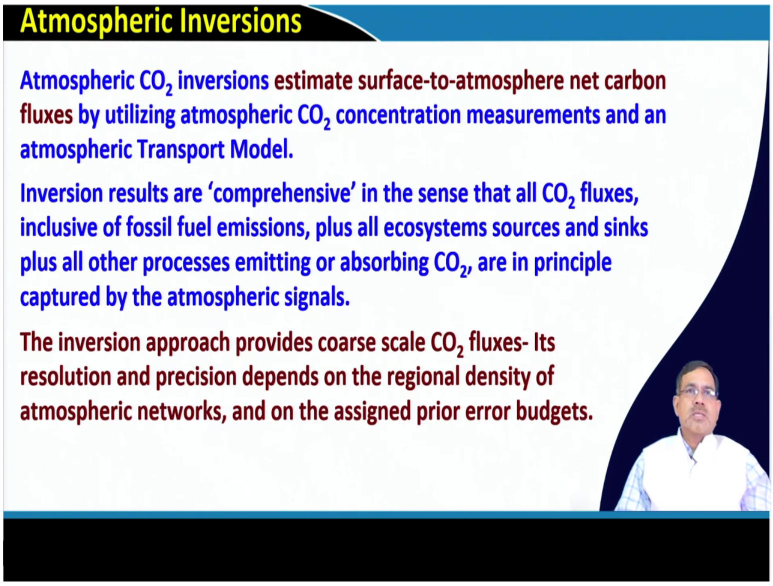
(Refer Slide Time: 20:49)



So, it could be vegetation inversion, it could be atmospheric inversion and in terms of atmospheric inversion, the different components of the inversion models are shown in this diagram taken from the Wageningen University group and these components could be in terms of observed atmospheric concentration of carbon dioxide, observe a sea surface concentration of CO 2 and a kind of general circulation model.

And believe me this is improving, but this is just a depiction because the contribution of transport model also is very important and we need to optimize it in the inversion.

(Refer Slide Time: 21:34)



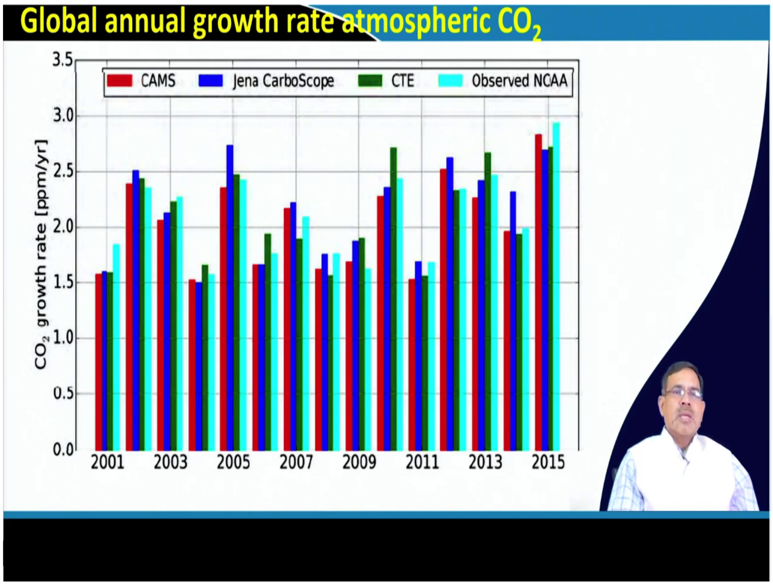
So, coming to this the key points; the atmospheric CO2 inversions estimate surface-to-atmosphere net carbon fluxes by utilizing atmospheric carbon dioxide concentration measurements and an atmospheric transport model.

So, inversion results are comprehensive in the sense that all carbon dioxide fluxes, inclusive of fossil fuel emissions, plus all ecosystem sources and sinks plus all other processes emitting or observing carbon dioxide are in principle captured by the atmospheric signal.

So, this we discussed. So, because this is a comprehensive because everything, all the sources of emission, all the sources of absorption are in principle captured by the atmospheric signal. So, that is the very important point as far as the atmospheric inversion is concerned.

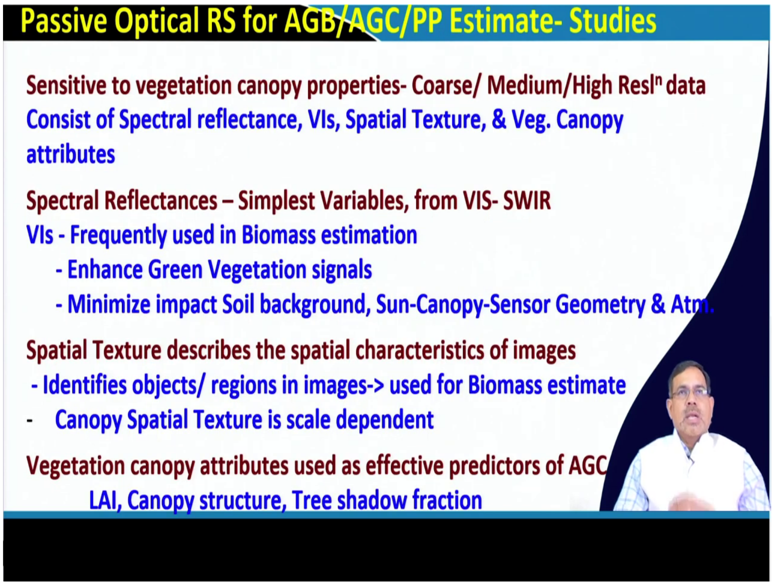
And then, the inversion approach provides coarse scale carbon dioxide fluxes, it could be in terms of kilometers by kilometers. So, its resolution and precision depends on the regional density and atmospheric network and on the assigned prior error budgets which are increasingly improving day by day, by different scientists and researchers. So, we must start understanding more and more in terms of atmospheric inversion and work more on this.

(Refer Slide Time: 23:03)



And coming from the three-model data with respect to the global annual growth rate in terms of atmospheric CO 2, the CAMS, Jena CarboScope, CTE, observed NOAA so, we see these are the different annual growth rates and in term and easy to understand.

(Refer Slide Time: 23:23)



So, now, we are done with all these methods in terms of the primary production estimates using the remote sensing. Now, let us try to wrap up the methods particularly the passive optical remote sensing methods for AGB or AGC or primary production so, above ground biomass, above ground carbon or primary production estimate studies.

So, friends as you know these the primary optical; the passive optical remote sensing is sensitive to vegetation canopy properties with respect to all the different resolution data available and consist and could consist of spectral reflectance, direct spectral reflectance bands, vegetation indices, spectral texture and vegetation canopy attributes.

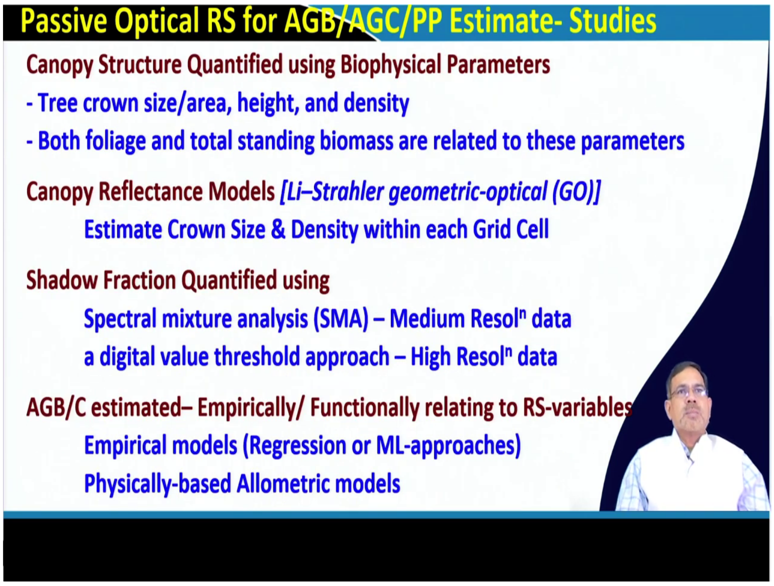
So, in terms of spectral reflectance, as we know they are very simple variables, or you can say the simplest variables and mostly range between visible and visible to SWIR range or to a infrared range as you know most of this vision VIs are sorry the reflectance bands in the visible to SWIR range gives us lot of information.

Coming to the VIS, the vegetation indices they are very frequently used in biomass or you say above ground biomass and above ground carbon estimation because they enhance the green vegetation signals are as we have been discussing. And all the VIs, they minimize the impact of soil background, sun-canopy-sensor geometry and atmosphere and space coming to the spatial texture, so, spectral reflectance, then vegetation indices, then the spatial texture.

So, spatial texture describes the spatial characteristics of images in terms of identifying the objects or regions in image and they are used for biomass estimate and as we know the canopy spatial texture is scale dependent when you use 3 by 3 window, when you use 5 by 5 window as you do, they will give different if.

So, texture the spatial texture means provides the spatial characteristics of image and that is why I use it for getting information on the above ground biomass or above ground carbon. And coming to the vegetation canopy attributes, they are also used for the AGC or above ground carbon estimate in terms of LAI what we know and also the canopy structure, tree shadow fraction.

(Refer Slide Time: 26:30)



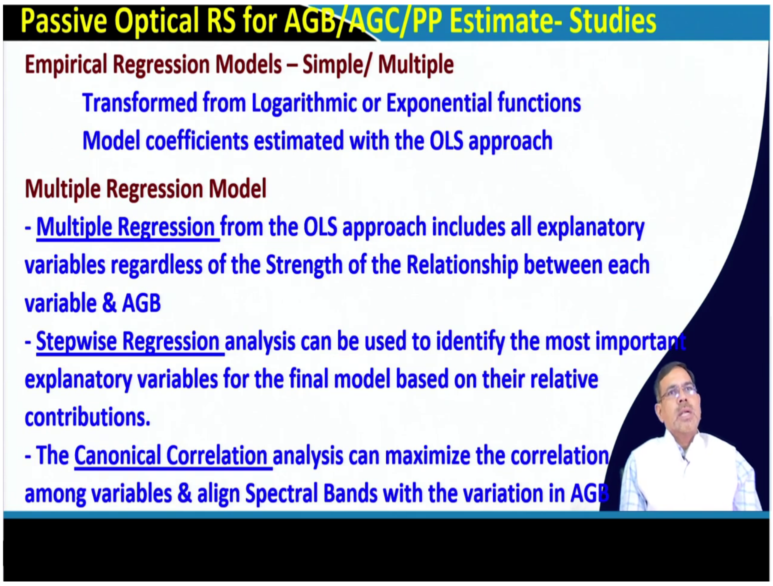
So, let us see in terms of the canopy structure quantified using biophysical parameters, it could be tree crown size, area, height, density, both foliage and total standing biomass are related to these parameters. In terms of the canopy reflectance model, they estimate the crown size and density within each grid cell.

So, coming to the canopy reflectance models, the canopy reflectance models estimate the crown size and density within each grid cell. One of the example is the Li-Strahler geometrical optical model GO.

Next is the shadow fraction, which is quantified, or which can be quantified using the medium resolution data and high-resolution data. So, medium resolution data we can do using spectral mixer analysis SMA whereas, for high resolution data, a kind of digital value thresholding approach does good.

We as far as the biomass or upper ground biomass and carbon estimation is concerned, it could be both empirical and functionally released related to our remote sensing based variables. So, empirical it could be regression based or machine learning approaches, physical attribute or physical based allometric models are also there.

(Refer Slide Time: 28:04)

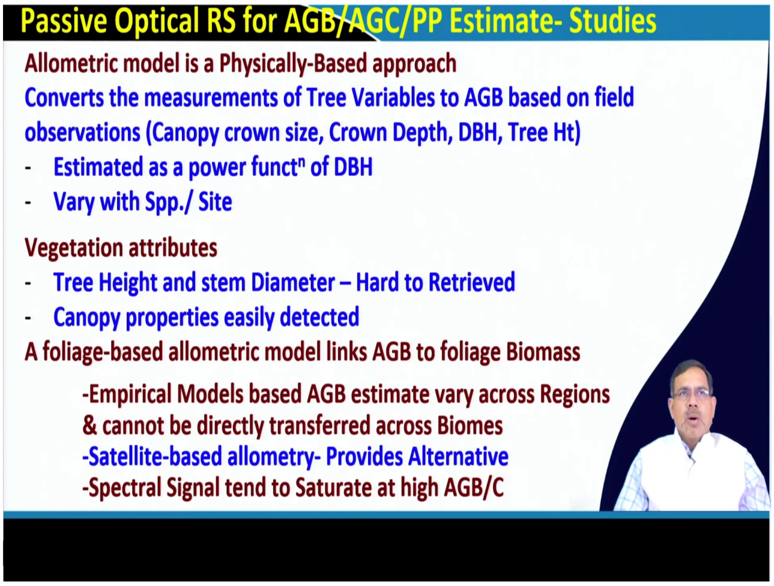


So, in terms of empirical models, it could be very simple models or multiple regression models. So, they in terms of simple, transferred from logarithmic or exponential function and the model coefficients estimated with the OLS approach. So, in terms of the multiple regression model, it could be multiple regression, stepwise regression or canonical correlation.

So, multiple regressions from OLS, ordinary least square approach includes all explanatory variables regardless of strength of the relationship between each variable and the above ground biomass whereas, the stepwise regression analysis can be used to identify the most important explanatory variable for the final model based on their contributions.

The canonical correlation analysis-based multiplication models can maximize the correlation among variables and align spectral bands with where the variation of the above ground biomass.

(Refer Slide Time: 29:16)



Coming to the challenges in terms of the allometric model, the allometry model is a physically based approach because it converts the measurements of tree variables to above ground based on above ground biomass based on the field observation; such as when we go to field, we observe lot of parameters and those parameters or those means field-based inputs are the key because to establish a allometric model.

So, canopy crown size, crown depth, DBH diameter at breast height what we measure on the field, the tree height which you measure using a hypsometer or a laser range finder or other equipments. So, in this allometric models, we estimate them particularly as a power function of DBH, the GPP or primary production estimate which also could vary with respect to species or the side to side.

So, vegetation attributes particularly the tree height and stem diameter are very hard or is difficult to retry from the passive optical remote sensing. So, in turn, in our next lecture, we will discuss this with respect to the optical remote sensing the SAR or microwave and lidar which also you know by now.

So, they are very useful in terms of giving the three-dimensional or the structural information like tree height and stem diameter which really holds lot of information with respect to the biomass and that is why the carbon and can be linked to the primary production.

And in terms of this passive optical remote sensing, the canopy properties ok; the canopy properties are very well are very easily detected. So, what happens if the passive optical remote sensing gives us a canopy property very easily or very well it is detected? So, the foliage-based so, canopy or foliage-based allometric models are the key to link the above ground biomass ok to link the biomass with the foliage biomass.

So, the allometry with respect to foliage and the AGB has are linked in a kind of allometric equation. So, this so, let me be very clear. So, from optical passive remote sensing data, you can get the canopy properties very well so, that could be linked in an allometric model to the above ground biomass ok.

So, in terms of some of the limitations yes, in the empirical models based AGB estimate vary across region and cannot be directly related or transferred across biomes because as we know these allometric based models are dependent on the field based and other enumeration based inputs and those are actually species based or location based. So, it is very difficult to generalize or transfer it across biomes. So, though they can at best hold good for that ecosystem.

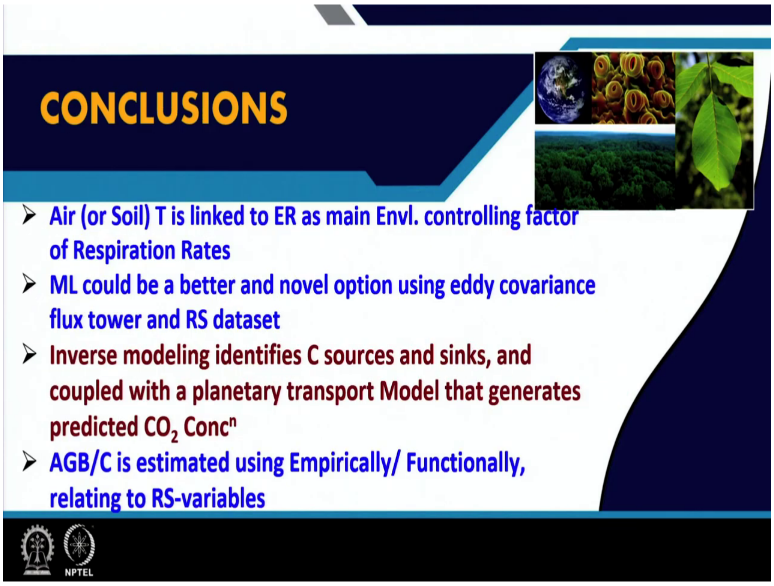
Second the satellite based allometry provides an alternative, it is just a linkage allometric linkage so, it is an alternative. However, the spectral signals tend to saturate at high AGB and or you say the high above ground carbon. So, when you are doing this in terms of allometric equation as we know at high biomass range which we mostly expect in terms of tropical forest so, the such the signal, the spectral signature from the passive optical remote sensing gets saturated.

So, there it is difficult if the above ground biomass level is very high, very high means mostly beyond 250 or 300 tons per hectare. So, that also we have to keep in mind and that leads to the next discussion with respect to our microwave or SAR and the lidar based utility.

(Refer Slide Time: 33:49)

So, coming to references yes, these are three articles we have been following in terms of our references.

(Refer Slide Time: 33:55)



And in terms of the conclusions, these are the same four points that have been mirrored here which are the same in (Refer Time: 34:02) in terms of the key points. So, we finish here the 3rd week discussion in terms of the primary product production.

So, we discussed what is primary production starting from the very basic concepts of primary production NPP, GPP, respiration, in terms of autotrophic and heterotrophic respiration, ecosystem production, biome production, ecosystem respiration and things like that.

So, with this, basic understanding we try to see that what, how the production or we say the primary production can be estimated using different methods. We have a very comprehensive idea that how they are estimated. We also have understood that what could be their possible limitations or uncertainties which will be elaborating in our in some of our future lectures in next week.

And also, we try to see some of the advance protocols or methodological protocols in terms of SIF solar induced fluorescence or model inversions. So, let me tell you the future is very bright and promising as far as the remote sensing based primary production studies are concerned because it gives us an estimate with respect to the vegetation-based carbon fluxes or that is what we call the primary production.

Thank you very much.