August 20- Phase 1: Conceptual Model Development and Data Compilation

This study aims to estimate future forest volume and stock in South Korea as affected by changes in environmental conditions (we will determine the model simulation period in this month).

The first task in this month is to compile existing scientific data, publications and information from Korea's National Forest Inventory and to develop a conceptual approach for the implementation of environmentally sensitive growth functions in the GCBM. First, I will review the relevant literature, search for the suitable study area for application of GCBM in South Korea, identify which data are available, which structure of data can be used, which spatial resolution is appropriate, which species to include and how to set the model simulation condition etc. for application of GCBM in South Korea.

The goal of the first period is to conceptualize the approach and to prepare a written document of the proposed approach for review by other experts. Throughout this conceptual development period, I will regularly communicate with Prof. Lee, Dr. Kim, Dr. Kurz and Max Fellows.

Conceptual approach and environmental variables: data availability, structure of data, spatial resolution etc.

GCBM is based on growth model to estimate and predict carbon balance of forest. Considering cycle of carbon pool in forest, stocks of carbon reservoir consist of five tree biomass components such as merchantable, other, foliage, coarse roots, fine roots etc.
Stand yield table is used to predict change of carbon balance for biomass, dead organic matter (DOM), soil carbon etc. of present and future. However, coefficients for growth model need to be estimated for the application of GCBM to forest in South Korea.
To run GCBM, some input data should be collected such as climatic data (mean temperature and precipitation); scenario for forest management (thinning and cutting); forest distribution; forest disaster.
Considering forest cover map and distractive district, stand level data is required and then stand volume is estimated through growth model reflecting pattern of forest growth.

Existing scientific data and environmental variables:

<table>
<tr><th>Data</th><th>Contents</th><th>Etc.</th></tr>
<tr><td>National Forest Inventory</td><td>Sample points of each 4 by 4 kilometers grid from field data
species , height, age, dbh, nha, etc.
(point)</td><td>Site index (SI)</td></tr>
<tr><td>Forest Cover Map</td><td>1: 5 000 scale forest type maps
Coniferous, broadleaf, mixed forest
Stand-level, Species, age, diameter class
(polygon)</td><td rowspan="2">In this study, a cadastral map was used to reflect land boundaries such as settlement, agriculture etc. Boundaries about forest are different with forest cover map and cadastral map.</td></tr>
<tr><td>Cadastral map</td><td>It is composed of a land map and a forestry map which is a continuous form of geographic information system.
(polygon)</td></tr>
<tr><td>Digital Elevation Model</td><td>Aspect, slope, elevation
(raster)</td><td>Topological wetness index (TWI)</td></tr>
<tr><td>Climatic data</td><td>Temperature, precipitation
Recent 10 years (2010 - 2019) climatic data from Automatic weather system (AWS) of 510 station in South Korea</td><td>Interpolation through kriging (precipitation) and IDW</td></tr>
</table>

	(point)	(temperature)

DBH and dominant tree height with each species are needed to notify forest growth and then site index could be estimated. Korea forest service provides yield table which contains 11 species such as Pinus densiflora, Pinus rigida, Pinus koraiensis, Larix kaempferi, Chamaecyparis obtuse, Quercus acutissima, Quercus variabilis, Quercus mongolica, Betula platyphylla, Liriodendron tulipifera. Although NFI deals with a lot of species in South Korea, 11 species will be used for this study due to insufficient data of yield table. Volume stock then will be estimated using DBH and height of tree (However, I have to check some coefficients for estimating SI and DBH etc. I'm checking it.)

Site index (SI) which expresses the site quality can be estimated through NFI. However, NFI consists of not all points covering Korea peninsula but sample points, about 4 kilometers. So, interpolation will be needed to acquire the intermediate data.

All data should have same resolution for correct analysis. So, spatial resolution is set as 1ha (100m x 100m) which is the lowest unit of compartment for forest management in South Korea. It was made for convenient management of forest. Most forest is managed by the compartment in South Korea.

National Geographic Information Institute (NGII) provides a digital elevation model (DEM) which consists of a high resolution of 5 meters. In this study, the data is resampled by resolution of 100 meters to set same resolution with other spatial data. Through it, TWI is calculated.

Korean meteorological administration (KMA) provides climatic data of mean temperature and precipitation in South Korea. Interpolation methods such as kriging and Inverse distance weighted (IDW) were used to estimate intermediate values of the independent data. In this study, temperature will be interpolated by using IDW and precipitation will be interpolated by using kriging.

Publications

Byun, J. G., Lee, W. K., Kim, M., Kwak, D. A., Kwak, H., Park, T., ... & Saborowski, J. (2013). Radial growth response of Pinus densiflora and Quercus spp. to topographic and climatic factors in South Korea. *Journal of Plant Ecology*, *6*(5), 380-392.

There are some data and variables such as tree age, DBH, TWI , climatic factors, A1B scenario, future climatic data (MM5).
Standard growth (SG) is defined as
$\Delta r = a \cdot age^{b}$,
Δr is annual radial growth, and a and b are coefficients of non-linear regression models.

To eliminate the influence of age on radial growth, standardization of radial growth with a fixed age, 30. SG eliminates the influence of age on tree growth while still incorporating impacts of competition, site quality and climate into Δr.
$$sg = \Delta r \cdot \left(\frac{30}{age}\right)^{b}$$

SG can compare under the same conditions. Relationship between annual radial growth,climate and topography can be analyzed without age. SG can have spatial autocorrelation in the spatial unit of a growth pattern.
Variogram is a general form for plotting the spatial autocorrelation of data which can be parameterized as range, nugget, sill values.

For Estimation of standard growth, GAM (Generalized Additive Model) is used for focusing on exploring data non-parametrically. The strength of the GAM is to deal with highly non-linear and non-monotonic relationships between independent and explanatory variables. GAM is especially useful for modelling species distributions because the smoothing functions can describe complicated non-linear relationships. GAM is for describing species response curves about environmental gradients, enabling prediction of species distributions under changing environmental variables.

eSG can be normalized with mean estimated SG (meSG) from integrating the normalized SG with the radial growth model. It could estimate the radial growth for variations because topographic and climatic conditions decrease. For fitting GAM, use of dataset within warmth index (WI) ensures the changing tendency of growth to temperature change. WI is the sum of values which are positive differences between temperature and 5℃.
$$eSG = f(TWI, T, P) = \beta_0 + \beta_1 \cdot TWI + \beta_2 \cdot T + \beta_3 \cdot P$$
$$\Delta\widehat{r}_{ie} = \widehat{r}_{i-5} \cdot \left(\frac{age_i}{age_{i-5}}\right)^{b} \cdot \frac{eSG}{meSG}$$

Sensitivity analysis compares changes in forest type according to differences in present and future growth using SG. Sensitivity to climate change is defined as the changing frequency of vegetation types in a grid area. The vegetation type in a certain area is likely to change if it is not suited to the new climate conditions.

Kim, M., Lee, W. K., Choi, G. M., Song, C., Lim, C. H., Moon, J., ... & Forsell, N. (2017). Modeling stand-level mortality based on maximum stem number and seasonal temperature. *Forest Ecology and Management*, *386*, 37-50.

The main objective of this study is to develop a model to simulate stand-level mortality for temperate forests in South Korea and to evaluate the effect of the climate factor such as temperature, on tree mortality. National Forest Inventory, Sterba's theory, semi-variogram analysis, and residual analysis were applied.

This study was processed for using 1484 permanent plots about five main temperate tree species of NFI. The stand self-thinning theory was extended to understocked stands which reflect that self-thinning starts before a stand reaches maximum density, and the self-thinning rate increases with stand density and reaches a maximum at the maximum stem number (MSN).

To develop a self-thinning model, ho, DBH, and SN were observed to fit the parameters of the MSN. ho is one of the most commonly used indicators of site productivity because a close correlation exists between volume and site index. It is generally accepted that the height of ho is minimally affected by competition.

For determination of the patterns of maximum stem number per ha about dominant tree height and development of a stand level mortality model using the MSN curve, and assessment of the impact of temperature on tree mortality in semi-variogram and linear regression models

The MSN curve expresses the upper boundary of observed stem numbers per ha. However, spatial autocorrelation was detected from residuals.
To analyze requirements nation-wide, data should deal with various variability such as forest structure, tree species, site. National Forest Inventory is the best for requirements of the data.

Four statistical methods were used for model validation of mortality function for temperate forest; Mean deviation, mean absolute deviation, standard deviation of difference, root mean square deviation (RMSD).
Tree mortality is a natural ecological process

To find the spatial autocorrelation if climate affects tree mortality, the spatial autocorrelation of residuals was checked about the difference between the observed value of the dependent variable and the predicted value.

Because tree mortality is relatied with climatic water and heat stress, relationship between the seasonal mean temperature (2006 - 2013) and residuals is identified using linear function.

Kim, M., Lee, W. K., Son, Y., Yoo, S., Choi, G. M., & Chung, D. J. (2017). Assessing the impacts of topographic and climatic factors on radial growth of major forest forming tree species of South Korea. *Forest ecology and management*, *404*, 269-279.

Tree ring growth plays an important role in identifying the growth response of trees to environmental and climatic variation. Some factors affecting the growth such as age, size of trees, climatic and topographic factors should be incorporated into the growth model.
The main goal of this study is to develop a tree- level radial growth model incorporating topographic and climatic factors for four major temperate tree species in South Korea. The standard growth model, semi-variogram analysis, and the generalized additive model (GAM) were used to achieve the goal.

Warmth index (WI) is appropriate for distribution of vegetation type and tree growth in South Korea and precipitation effectiveness index (PEI) is useful to classify climatic zones corresponding to forest ecosystem types and to demonstrate the relationship between a hydrological index and vegetation regime. In this study, summer PEI (June to August) is used instead of annual PEI because some species such as red pine, cork oak, Mongolian oak, and Korean pine is positively correlated with the annual radial growth. Topographic wetness index (TWI) is the most popular indicator to be correlated with soil attributes including horizon depth, silt percentage, and organic matter as a predictor variable for forest health conditions.

Two approaches to evaluate diameter growth are to use core samples from the NFI and standard growth (SG) model.

$$MR_{ij} = \sum_{k=1996}^{2005} AR_{ijk}(mm) \div 10$$

SG model is from power function which is used to extract the general age-related growth patterns of each species. In this study, SG was defined as the radial growth at 30 years. To convert the MR from various tree ages to the SG at age 30, the transformation to algebraic differences form was applied. Individual trees of different age can be compared under equivalent conditions by eliminating the effect of tree age on tree growth, which can analyze the relationship between SG, climate, and topography, without age dependent responses, quantitatively.

$$SG_{ij} = MR_{ij} \cdot \left(\frac{30}{age_{ij}}\right)^b$$

Semi-variogram analysis is used to identify spatial autocorrelation fitted to the spherical model.
Generalized additive model (GAM) with a spline function is used to analyze models. The GAM is a nonparametric extension of the generalized linear model (GLM).

$$SG_{ij} = \beta_1 \cdot WI_i^2 + \beta_2 \cdot WI_i + \beta_3 \cdot PEI_i^2 + \beta_4 \cdot PEI_i + \beta_5 \cdot TWI_i + Int$$

Optimal ranges of climatic conditions for growth are set for optimal range for each species. Climate factors and tree growth were modeled using a non-linear quadratic function except for the relationship with TWI, which was modeled as a linear function based on previous research.

SG is defined as the expected radial growth at 30 years of age given the environmental conditions at the tree's location. Thus, eSG can indicate the relative spatial suitability of the different species. Alternatively, the eSGs of one species can indicate the relative suitabilities of the environmental conditions at different sites. In conclusion, eSG can be used as an independent variable in the growth model. The result of the eSG for each tree species was indirectly validated using standardized indices of tree-ring width. The indices were produced using the C-method. The C-method is a tree-ring standardization method based on the assumption that the annual growth rate of mature trees fluctuates around a specific level, expressed by a constant ring width. Of the standardization methods based on the biological age tree rings, the C-method has the advantage of calculating an expected growth curve for each measurement series, whereas the regional curve standardization applies the same growth curve to all samples. In this study, the radial growth model was developed to account for the effects of climatic and topographical factors on diameter growth using eSG.

$$\Delta\widehat{r}_{fij} = \Delta\widehat{r}_{pij} \cdot \left(\frac{age_{ij}}{age_{pij}}\right)^{b} \cdot \frac{eSG_{fi}}{meSG_{p}}$$

Piao, D., Kim, M., Choi, G. M., Moon, J., Yu, H., Lee, W. K., ... & Cui, G. (2018). Development of an Integrated DBH Estimation Model Based on Stand and Climatic Conditions. *Forests*, *9*(3), 155.

To analyze climate and to identify characteristics of diameter development for major species, DBH estimation model was developed based on age, site quality and stand density.

NFI data was divided with two parts: model development (90%) and validation (10%).

In addition, tree growth was classified into three spatial categories: stand-, watershed-, and regional level. Watershed level factors such as aspect and slope do not have spatial autocorrelation though elevation, TWI, temperature and precipitation show spatial autocorrelation at ranges over 4 kilometers.

In analysis of regression model, integration of non-spatial and spatial data might lead to low significance. So, this study was performed by three steps. First, development of a diameter model for DBH estimation using stand-level factors without spatial autocorrelation. Second, development of a residual model to predict the relationship between residuals climate factors with spatial autocorrelation including temperature and precipitation. Third, integration of the diameter model and residual model to develop an integrated DBH estimation model, in which DBH could be predicted by climate factors and stand-level factors.

Some factors such as age, SI and Nha were used to develop DBH estimation model. Nonlinear regression was performed because of better performance than linear regression of DBH development in a stand.

Semi-Variogram introduces deviation from the independent-observation assumption of classical statistics. If difference in residuals exists at the regional level by climate and topography, residuals will show spatial autocorrelation. The relationship between climatic factors and residuals of DBH was modeled using a simple linear function and then DBH model was derived by combination using results of the linear regression analysis.

Others

GAM

The terms $f_1(x_1),\cdots,f_p(x_p)$ denote smooth, nonparametric functions, which mean that the shape of predictor functions are fully determined by the data as opposed to parametric functions that are defined by a small set of parameters.

Spatial autocorrelation

Spatial autocorrelation is analyzed by the ratio of the nugget to the sill. The sill means the maximum observed variability of the data. And the nugget means the variability of the field data which cannot be explained by the distance among the observations. The chance of spatial autocorrelation is low when the ratio is high. Adversely, ratio is low, there is spatial autocorrelation

Spherical model (semi-variogram)

Spherical model is the most commonly used variogram model. It is characterized by a very steep, exponential increase in semi-variance. That means it approaches the sill quite quickly. It can be used when observations show strong dependency on short distance.