Landslide Susceptibility Analysis by Terrain and Vegetation Attributes Derived from Pre-event LiDAR data: a case study of granitic mountain slopes in Hofu, Japan

Junko IWAHASHI*, Takaki OKATANI, Takayuki NAKANO, Mamoru KOARAI, and Kosei OTOI

Geospatial Information Authority of Japan (Kitasato-1, Tsukuba, Ibaraki 3050811, Japan)
*Corresponding author. E-mail: iwahashi@gsi.go.jp

This study explores a method of creating tree height data and the root strength index data using archived LiDAR data, which include coarse and leafless season data, and improves the assessment of susceptibility of granitic mountain slopes to rainfall-induced landslides, considering vegetation in addition to topography. The study areas are located in mountain forests in central and western Japan. We found that the tree heights of broadleaf deciduous forest estimated using DCM (Digital Canopy Model) should be corrected by trigonometry, and the average values of the DCM tree heights in the 30-m grid could be verified even in broadleaf deciduous forests in comparison with the field data. We proposed a new factor “the root strength index,” which is the product of the tree height and the square root of the tree density. In the Hofu region, the root strength index estimated from the pre-event DCM is inversely proportional to the rate of rainfall-induced landslides occurred in July 2009. The use of the root strength index in addition to topographic attributes partially improves the correct prediction rate of rainfall-induced landslides in the Hofu region. Concave slopes, and strongly concave steep slopes in particular, show clear improvement in the correct prediction rate through the use of the root strength index.

Key words: landslide, LiDAR, digital canopy model, tree height, the root strength index

1. INTRODUCTION

The GSI (Geospatial Information Authority of Japan) archives contain wide ranges of LiDAR (Light Detection and Ranging) data measured by subordinate agencies of MLIT (Ministry of Land, Infrastructure, Transport and Tourism) as well as the GSI, which can be used for terrain surveys and disaster prevention.

Shallow landslides were investigated in a previous study using two topographic attributes, the slope gradient and the convexo-concave index (the Laplacian) calculated from DEMs [Iwahashi et al., 2012]. In that study, the authors found that the representative window sizes are approximately 30 m for rainfall-induced shallow landslides, and the optimal window size may be directly related to the average size of landslides in each region. The authors also found a stark contrast between rainfall- and earthquake-induced landslides. Rainfall-induced landslides are most commonly observed to occur at a slope gradient of 30°, and at a convexo-concave index of valley heads. The spatial distribution of shallow landslides in Tertiary sedimentary rocks and the influences of stratal architectures and artificial changes using the data of repeated landslide events and LiDAR DEM have also been investigated [Iwahashi and Yamagishi, 2010]. However, the effect of vegetation was not considered in those studies.

It was revealed that tree roots improved the stability of hillslopes [Waldron, 1977; Abe and Ziemer, 1991]. The soil binding power of tree roots increases relative to the trunk diameter [Abe, 1997; Yamaba and Sano, 2008], and trunk diameter is proportional to tree height [Shimada, 2011]. Tree height indirectly suggests the soil binding power of the tree roots. If the soil binding power is considered in forests, the tree density must be useful information. We therefore address the tree height and tree density as the index of the soil binding power of tree roots.
This study has two challenges, of which the first is data creation. In the field of forestry, many researchers [e.g., Clark et al., 2004; Itoh et al., 2009] reported that DCMs (Digital Canopy Model: difference of DSM and DEM) correspond with the actual tree heights. However, those studies used very high resolution data of evergreen forests. MLIT has obtained wide ranges of LiDAR data, although densities of archived data are often coarse (about 1 to 3 pt/m² in the early 2000s), and the data were measured mainly in leafless season for conducting terrain surveys. Therefore, the tree heights from the archived LiDAR data should be compared with ground truth data, especially in the forests other than evergreen forests. The second challenge for this study is landslide assessment. It is well known that occurrence of rainfall-induced shallow landslides increase due to deforestation [e.g., Glade, 2003]. However, a quantitative relationship between the tree heights and the frequency of occurrence of rainfall-induced shallow landslides remains to be understood. In this study, we compared pre-event vegetation data with the landslide inventory map. This study sets a goal of developing a method for assessment of rainfall-induced shallow landslides using vegetation data in addition to topographic attributes.

2. STUDY AREA AND TREE MEASUREMENTS IN THE FIELDS

The data creation challenge is explored in five regions of Japan (Fig. 1). Table 1 describes the study areas. We set six to eight measurement fields per region, 33 fields in total, for conducting tree measurements. The size of measurement fields were 20 m × 40 m (21 fields), 20 m × 50 m (nine fields), 20 m × 20 m (two fields), and 20 m × 80 m (one field), which are close to the 30-m grid in area. Approximately 3,000 trees were measured in the 33 fields. Locations of the measurement fields were traverse surveyed by a total station from identified RTK-GNSS points. Tree locations in the fields were determined by compass surveying. Tree heights were measured using ultrasound measurement instrument (Vertex IV, Haglof Inc.). Vertical measuring error using Vertex VI is defined within 10 cm, though measuring results could have more errors by human-incident. Databases of tree height, tree species, trunk diameter, and tree numbers in the measurement field were created. The GIS data of trunk positions and canopy polygons were also created. The tree heights correlated well with trunk diameters.

Table 1 Summary of the measurement fields.

<table>
<thead>
<tr>
<th>Region</th>
<th>Month and year of tree measurement</th>
<th>Main species*</th>
<th>Main lithology of bed rocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hofu</td>
<td>Sep. to Nov. 2012</td>
<td>BDF 2, NEF 2, BEF 2</td>
<td>Cretaceous granite</td>
</tr>
<tr>
<td>Izumozaki</td>
<td>Oct. 2011</td>
<td>BDF 3, NEF 4, Mix 1</td>
<td>Tertiary sedimentary rocks</td>
</tr>
<tr>
<td>Niihama</td>
<td>Sep. to Nov. 2012</td>
<td>BDF 2, NEF 2, Mix 1</td>
<td>Cretaceous sedimentary rocks</td>
</tr>
<tr>
<td>Shobara</td>
<td>Sep. to Nov. 2012</td>
<td>BDF 3, NEF 4</td>
<td>Cretaceous rhyolite</td>
</tr>
<tr>
<td>Aso-Ichinomiya</td>
<td>Aug. 2013</td>
<td>BDF 1, NEF 6</td>
<td>Quaternary volcanic rocks and pyroclastic deposits</td>
</tr>
</tbody>
</table>

The landslide assessment was conducted in the Hofu region (47 km² around Hofu City, Yamaguchi Prefecture, Japan; Romanized as “Houfu” in Iwahashi et al., 2012) where large numbers of shallow landslides occurred due to heavy rainfall in July 2009. The pre-event LiDAR data measured in 2005 covers the damaged area widely. The lithology of the Hofu region is mostly comprised of Late Cretaceous granite. Intrusive or metamorphic rocks and alluvium are distributed more rarely [Matsuura et al., 2007]. The Hofu region lies on moderate mountains under altitudes of 500 m. The July 2009 heavy rainfall event around the Hofu region was...
characterized by torrential rains in the rainy season, with daily precipitation reaching 275 mm [Misumi, 2010]. Over 1,000 landslides and subsequent debris flows occurred in deeply decomposed granite, and 17 victims died including those affected by a landslide that occurred in an upstream slope of a nursing home [Misumi, 2010].

3. VEGETATION ATTRIBUTES

3.1 LiDAR Data

Table 2 shows the specifications of LiDAR data that were used in this study. We obtained DCM (Digital Canopy Model) from the difference in DSM (Digital Surface Model) and DEM (Digital Elevation Model). The DEMs used were outsourcing products. We generated DSMs from point clouds of original LiDAR data. DSM data were generated by taking maximum heights of point cloud for each rasterization grid.

Table 2 Summary of the LiDAR data.

<table>
<thead>
<tr>
<th>Region</th>
<th>Month and year of survey</th>
<th>Density* (pt/m²)</th>
<th>Condition of the leaves of deciduous tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Izumozaki</td>
<td>Nov. 2007</td>
<td>3.0</td>
<td>Half fallen</td>
</tr>
<tr>
<td>Niihama</td>
<td>Dec. 2008</td>
<td>1.8</td>
<td>Mostly fallen</td>
</tr>
<tr>
<td>Shobara</td>
<td>Mar. 2012</td>
<td>27.5</td>
<td>Totally fallen</td>
</tr>
<tr>
<td>Aso-Ichinomiya</td>
<td>Jan. 2013</td>
<td>8.5</td>
<td>Totally fallen</td>
</tr>
</tbody>
</table>

3.2 Tree Height

It was often difficult to estimate individual tree height from LiDAR DCM, because the densities of point clouds of the archived data were frequently insufficient. In addition, several kinds of trees are tilting, and tree crowns overlap one another. The field data of tree height may thus include survey error for the location of tree crown and tree heights. We assumed that the uppermost canopy polygons in the GIS data include tree tops. The uppermost canopy polygons of large trees (trunk diameter ≥ 10 cm) were extracted, and the field measured data of their tree heights were compared with DCM. Some of the LiDAR data had been measured for several years before the field measurement of tree heights was conducted (Tables 1, 2). DCM values had been corrected according to the estimated grown-up heights using forest tree growth curves that were published by local governments for forestry. We assumed a moderate site index and corrected older DCM data using the curves of prefectural or neighboring prefectural area. The curves of oak were used for broadleaf deciduous forests. Distributions of broadleaf deciduous forests had been estimated from the results of the National Surveys on the Natural Environment (Biodiversity Center of Japan; GIS data of actual vegetation: http://www.biodic.go.jp/trialSystem/top_en.html).

Fig. 2 Average tree height for each measurement field.

One meter DCMs were corrected using only grown-up values.

Fig. 3 Average tree height for each measurement field.

One meter DCMs were corrected with grown-up values, and the broadleaf deciduous tree areas were corrected using a trigonometric function.

Fig. 2 compares the average DCM heights (corrected with grown-up values only) and the
average field measured tree heights in the measurement fields. In the case of broadleaf deciduous forests, the average DCM heights were overestimated due to tilting of trees. Evergreen trees tend to grow vertically, and broadleaf deciduous trees tend to grow perpendicular to slopes.

Therefore, we used the cosine of DCM height as the estimated tree heights of broadleaf deciduous forests and mixed forests. Fig. 3 shows the corrected values.

Fig. 3 shows the corrected average values for each measurement field. However, scatter grams of individual trees often show poor correlation, especially in the case of broadleaf deciduous forests.

Fig. 4 shows scatter grams of individual tree heights of broadleaf deciduous forests in the Izumozaki region. Red crosses indicate the averages that correspond to the dots in Fig. 3. Fig. 4 indicates that averages in the 30-m grid, which is nearly similar in area to measurement fields, can present good results even in broadleaf deciduous forests, where DCM values show almost no correlation with the field-measured individual tree heights.

Point cloud density of LiDAR data may have effects on the correlations of individual tree height with correspondent DCM height. In the Hofu region, the 1-m DCM created from the 2009 LiDAR data (11.2 pt/m²) shows good correlation (R² = 0.84) with the all field measured tree height of large trees (trunk diameter ≥ 10 cm). The correlation obtained using the 2005 LiDAR data (1.2 pt/m²) is lower. In addition, extremely fine resolution of point cloud density caused lower correlation (R² = 0.76 for 1-m DCM) than coarser but suitable resolution (R² = 0.79 for 2-m DCM). However, average tree heights of the 2005 DCM in measurement fields are not inferior in comparison with the 2009 DCM.

3.3 Tree Density

We calculated the tree density using the image processing method of Okatani et al. (2013). The method extracts cells that have the maximum DCM values within 3 × 3 cells, i.e., those cells that are estimated to be tree tops (Fig. 5) using the maximum filter. We defined the tree density as the numbers of the tops of tall (corrected DCM ≥ 5 m) trees, excluding shrubs, in the 30-m grid. This method can be applied to a wide range of conditions, including deciduous forests in winter, although the number of peaks may be influenced by DCM resolution.

Fig. 6 compares the tree densities estimated from 1-m DCM and measured data in the fields. Over the resolution limit, the estimated tree densities of needleleaf evergreen forests reach the maximum. However, the two data for leafy forests under the resolution limit show correlations. In addition, underestimated needleleaf evergreen forests, which are dense artificial cedar forests, include many thin trees or young trees according to the tree measurements. Therefore, very high tree densities do not indicate sufficient soil binding power of tree roots. We consider that the estimated tree density from DCM peaks could be supplementarily used for landslide assessment.
3.4 The Root Strength Index

The soil binding power of tree roots is proportional to the value of the trunk diameter divided by the distance between the trees [Tochimoto et al., 2010]. The trunk diameter is proportional to tree heights. The square of tree distance is inversely proportional to the tree density. Therefore, we defined the root strength index using LiDAR data as the product of the estimated tree height and the square root of estimated tree density, as Eq.(1).

\[ RST = H \times \sqrt{D} \]  

(1)

where RST: the root strength index, H: estimated tree height, D: estimated tree density.

The tree density may include some uncertainty about contribution to the soil binding power. Planting tree too thick can result in poor growth of tree roots, especially in artificial forests. However, the threshold of root growth in relation to tree density is not clear. In this study, that issue is not taken into account in Eq.(1).

4. LANDSLIDE SUSCEPTIBILITY AND VEGETATION

In this chapter, the authors introduce the results of landslide susceptibility analyses conducted using tree height and root strength index extracted from the pre-event (2005) LiDAR data of the Hofu region. The landslide inventory data were derived from 20 to 60 cm orthoimagerys obtained in August 2009.

4.1 Correlation between Landslide Susceptibility and Tree Height or Root Strength Index

We compared the estimated tree height from pre-event DCM and the rate of shallow landslides that occurred in July 2009 heavy rainfalls with the eliminating influence of topography (Fig. 7). The horizontal axis corresponds to the topographic vulnerability of the Hofu region [Iwahashi et al., 2012] calculated from slope gradient and the Laplacian. The topographic vulnerability caused due to the frequencies of landslide cells versus slope gradient or the Laplacian derived from the LOG filter (Laplacian of Gaussian; Marr and Hildreth, 1980), and their contribution, respectively. The two terrain attributes were calculated from the pre-event 2-m DEM in 15 × 15 cells (30-m window size). Then the topographic vulnerability data derived from the 2005 LiDAR DEM and the 2009 landslide inventory data were summarized in the 30-m grid. In Fig. 7, the depth axis corresponds to the tree height calculated from the 2005 LiDAR DCM and averaged in the 30-m grid. The Percentage of the 2009 landslide in Fig. 7 shows the percentage of 30-m blocks which include one or more 2-m landslide cell. Fig. 7 reveals that higher average tree heights tend to cause a lower rate of landslides even if the topographical vulnerability remains constant.

Fig. 6 The tree densities estimated from 1-m DCM and measured data in the fields.

![Graph showing tree density vs field data](image)

Fig. 7 Percentage of the 2009 landslides in the Hofu region, the topographic vulnerability, and the tree height derived from the 2005 LiDAR DCM. The data were summarized in 30-m grid, and the data sections less than 100 cells were omitted.

![Graph showing correlation between tree height and landslide rate](image)

Fig. 8 Percentage of the landslides that occurred during the 2009 heavy rainfall in the Hofu region, in comparison with the root strength index calculated from 2005 LiDAR DCM.

Fig. 8 shows the root strength index calculated from the pre-event (2005) 2-m DCM in the horizontal axis and the rate of landslides in July
2009 in the vertical axis, respectively. The figure reveals a clear correlation between the increase of the root strength index and the decrease of the rate of landslides.

4.2 Contribution of the Root Strength Index to Landslide Prediction
We estimated the contribution of the root strength index to landslide susceptibility in the Hofu region. The primary factor dataset was created from the pre-event 2-m LiDAR data. The primary factor dataset includes landslide occurrence as a categorical valuable; moreover, the slope gradient, the convexo-concave index (the Laplacian) described in section 4.1, and the root strength index were derived from DCM as continuous variables.

At first, we examined overall correct prediction rate using SVM (Support Vector Machine; Cortes and Vapnik, 1995). In this analysis, we used a 30-m grid summarized dataset considering the data-handling capacity. SVM is a supervised learning model that is highly applicable to non-linear data such as the rate of landslide, which is often non-linear to primary factors. SVM is an excellent classifier for two categories. We used the ksvm command of the kernlab library [Karatzoglou et al., 2013], which works in free-software R [R Core Team, 2013]. The 30-m grid landslide inventory data of the July 2009 heavy rainfall was coupled with the primary factor data, and non-failure cells were extracted into the same number of the failure cells using a random sampling technique. Then, the dataset was divided into two groups using a random sampling procedure. One group was used as training data and the other was used as prediction data. Consequently, overall correct prediction rate was determined to be 73.7% using the two topographic attributes (slope gradient and convexo-concave index), and 74.4% using the three primary factors, including the root strength index as well as the topographic attributes. Consequently, the total increase in correct prediction rate is very small even in the Hofu region where the correlation between landslide and vegetation is very clear (Figs. 7, 8). This indicates that the prediction rate is partially raised by using the root strength index.

Therefore, we then used original 2-m datasets and analyzed them after dividing the datasets into groups according to types of topography. In the case of rainfall-induced landslides such as the 2009 landslides in the Hofu region, the rate of shallow landslide is the highest around 30 degrees for slope gradient, and the value around valley head for the convexo-concave index [Iwahashi et al., 2012]. In addition, the convexo-concave index divides the convex and concave slopes at the zero value. We used the obtained threshold values to divide slopes into groups. Then, groups that include more than 1,000 failure cells were analyzed. A thousand each of failure and non-failure cells were extracted. We examined two methods, the SVM and a popular method, the linear discriminant analysis (LDA). In the SVM analysis, one group was used as training data and the other was used as prediction data. In the LDA, we compared the two classification results simply. The results are described in Table 3. Strongly concave slopes, steep strongly concave slopes in particular, show clear improvement in the correct prediction rate or the correct answer rate.

Table 3 The correct prediction rate of the 2009 landslides by SVM and the correct answer rate by LDA for each slope type. SG: Slope gradient (degrees). SVM: Support Vector Machine, LDA: Linear Discriminant Analysis.

<table>
<thead>
<tr>
<th>SVM</th>
<th>5 ≤ SG ≤ 30</th>
<th>30 ≤ SG ≤ 47</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly concave (Valley bottom to valley head)</td>
<td>56.9% ⇒ 60.9% (+4%)</td>
<td>55.5% ⇒ 63.9% (+8.4%)</td>
</tr>
<tr>
<td>Weakly concave (Valley head to ridge)</td>
<td>65.2% ⇒ 66.2% (+1%)</td>
<td>57.1% ⇒ 57.9% (+0.8%)</td>
</tr>
<tr>
<td>Convex slope</td>
<td>74.3% ⇒ 74.8% (+0.5%)</td>
<td>66.9% ⇒ 65.3% (-1.6%)</td>
</tr>
<tr>
<td>LDA</td>
<td>5 ≤ SG ≤ 30</td>
<td>30 ≤ SG ≤ 47</td>
</tr>
<tr>
<td>Strongly concave (Valley bottom to valley head)</td>
<td>59.3% ⇒ 61.8% (+2.5%)</td>
<td>51.5% ⇒ 63.3% (+11.8%)</td>
</tr>
<tr>
<td>Weakly concave (Valley head to ridge)</td>
<td>65.8% ⇒ 67.1% (+1.3%)</td>
<td>59.4% ⇒ 61.9% (+2.5%)</td>
</tr>
<tr>
<td>Convex slope</td>
<td>73.9% ⇒ 74.3% (+0.4%)</td>
<td>66.2% ⇒ 67.3% (+1.1%)</td>
</tr>
</tbody>
</table>

5. DISCUSSION
There are two factors that may have a correlation with the soil binding power of tree roots other than trunk diameter.

The soil binding power of tree roots may differ with tree species. According to research in Japan, oaks have stronger roots [Tochimoto et al., 2010]. However, deciduous trees in Japanese mountains, which are populated with a large variety of species including oaks, do not always express stronger tree roots than needleleaf evergreen trees, which are mostly Japanese cedar or cypress plantations [Kurokawa, 2012]. The root strength index of the Hofu region does not consider tree species, although Fig. 7 expresses a clear correlation. Difference of
the soil binding power in tree species should be offset by including diverse species in study area. Therefore, we consider that the root strength index calculated from LiDAR data is a new suitable index for landslide susceptibility assessment in regions where the influence of tree roots can be found.

The second factor is surface soil in relation to tree roots. Although occurrence frequencies of landslides in intrusive or metamorphic rock slopes are rare, granitic slopes in the Hofu region experienced large number and density of landslides in July 2009. This indicates differences in topsoils between deeply decomposed granite slopes and hard rock slopes. Such slopes may differ in tree root growth in bedrocks. This diversity of topsoils may be observed in the slopes of the same legend of the geological map. However, the method for determining the thickness of topsoil over a wide area, which is a key factor in shallow landslides, is currently not available. Under these circumstances, the root strength index derived from DCM is expected to improve landslide susceptibility assessment in concave steep slopes, which are understood to have high topsoil thickness. Landslides in strongly concave slopes were difficult to predict using the topographic attributes only; however, the use of the root strength index improved the rates (Table 3).

A practical effect of using the root strength index is in extracting a basin with high priority of landslide assessment. The black polygons in Fig. 9 represent the July 2009 landslides and their extents in a part of the Hofu region. The pink cells of Fig. 9-a represent the granitic slopes of 15 to 45 degrees in middle valley to ridge, which are highly susceptible to rainfall-induced landslides in general and also in the Hofu region [Iwahashi et al., 2012], from pre-event LiDAR DEM and the geological map [Matsuura et al., 2007]. Fig. 9-a indicates that topographic attributes extract a large number of object areas. Fig. 9-b adds the root strength index to the object areas of Fig. 9-a in color gradation. Although it is confined to a certain region where the influence of tree roots can be found, and to small shallow landslides that are not affected by stratal architectures of bedrocks, there is an obvious advantage in adding the root strength index to landslide assessment. Enhanced practical assessment of shallow landslides can be made possible by emphasizing use of the root strength index for more susceptible basins.

The expected ripple effects from this study include not only enhancement of the assessment of shallow landslides but also forest management and biomass assessment [Drake et al., 2002] by addressing the challenge of creating tree height data from the archived LiDAR data. Even if the influence of tree roots on landslides is minimal in this region, tree heights data are useful, for example, in order to obtain a rough estimate of the volume of trees displaced due to floods.

The remaining issue with this method is the creation of tree density data. However, the densities of archived LiDAR point cloud data continue to increase. The problem caused by the low density of point cloud could be resolved with time.

6. CONCLUSIONS

We have designed a method for calculating tree height, tree density, and the root strength index from the archived LiDAR data, which include coarse and leafless data in addition to field-surveyed data from five regions in Japan. We confirmed that the tree height and the root strength index had negative correlation to the rate of rainfall-induced landslides from a case study in the Hofu region that is characterized by granite mountains.

We compared the correct prediction rates when using only the two topographic attributes (slope
gradient and convexo-concave index) and in another case, using the three primary factors, which include the root strength index besides the topographic attributes. Total increase in the correct prediction rate considering the three primary factors was very small even in the Hofu region where the correlation between landslide occurrence and vegetation is very clear. However, concave slopes, strongly concave steep slopes in particular, show clear improvement in the correct prediction rate using the root strength index. Therefore in the case study, addition of the root strength index improved the prediction of shallow landslides.

This study may improve the assessment of shallow landslide, forest management, and biomass assessment.

ACKNOWLEDGMENTS: We are grateful to Dr. Izumi Kamiya of GSI for his helpful advice on statistical analysis. We are grateful to Dr. Hiromu Daimaru of Forestry and Forest Products Research Institute for his comments. We also thank two anonymous referees for their helpful reviews.

REFERENCES