3D imaging and quantitative analysis of peach tree architecture via *TreeQSM*

J. Knapp-Wilson[1,2], R. Bohn Reckziegel[3], A. Bucksch[4,5,6] and D.J. Chavez[1,2]

[1]University of Georgia, Department of Horticulture, Griffin Campus, Griffin, GA 30223, USA; [2]University of Georgia, Institute of Plant Breeding Genetics and Genomics, Griffin Campus, Griffin, GA 30223, USA; [3]University of Freiburg, Chair of Forest Growth and Dendroecology, Freiburg, 79098, Germany; [4]University of Georgia, Department of Plant Biology, Athens Campus, Athens, GA 30602, USA; [5]University of Georgia, Warnell School of Forestry and Natural Resources, Athens Campus, Athens, GA 30602, USA; [6]University of Georgia, Institute of Bioinformatics, Athens Campus, Athens, GA 30602, USA.

Abstract

Peach tree morphology and vigor has seen renewed interest since development of new semi-dwarfing rootstocks. Even so, fruit quality, yield, and disease resistance remain the primary traits of interest in most fruit breeding programs. As such, canopy morphology and tree architecture remain relatively untapped areas of improvement. These traits are considered challenging to study and phenotype because of their inherent complexity and quantitative nature. While challenging, devising methodologies that can quantify canopy morphology and tree architecture would greatly aid in alleviating agronomic burdens felt by growers and researchers. Better characterizing tree architecture would allow for easier identification of superior cultivars/genotypes that would innately require less pruning and/or training. This in turn would optimize resource utilization. To accomplish this, large amounts of branching data will be required to sufficiently study and quantify tree architecture. Traditional means of collecting branching data however are difficult. Most traditional methods are destructive and/or require manual counting/recording. Manually collecting branching data are labor-intensive, repetitive, and prone to human error. A more modern and novel approach to collecting this data are via 3D terrestrial laser scanning (TLS) technology, such as terrestrial LiDAR (tLiDAR). 3D tLiDAR scanners can generate point clouds of scanned trees that can be virtually modeled. Running multiple iterations of these modeling simulations should yield the necessary branching data to begin better characterizing tree architecture. Our goal was to test these 3D reconstructive models and assess their overall fit when compared to our scanned data. The field data, alongside the general fit of these models, provided clarity as to the reliability of the quantitative data recovered from our 3D scans/reconstructions. As a result, further studies into architecture and morphology will be made demonstrably more feasible with the provided methods.

Keywords: computational biology, phenomics, tLiDAR, horticulture, bioinformatics

INTRODUCTION

Peaches [*Prunus persica* (L.) Batsch] are a deciduous fruit tree grown in temperate regions across the world. In terms of production, peaches rank second only behind apples in terms of worldwide production, which has increased from 13.2 million tonnes in 2000 to 24.6 million tonnes in 2020, nearly doubling in size. This increase was also seen for the amount of land used for peach cultivation, growing from 1.27 million hectares in 2000 to 1.49 million hectares in 2020 (FAOSTAT, 2020). A majority of this growth occurred in China, the world's top peach producer. In 2000, China produced 3.8 million tonnes of peaches across 467,000 ha. This figure has grown massively, with China producing approx. Fifteen million tonnes of peaches across 780,000 ha. However, the United States peach production and land-use area has experienced a sharp decrease during the same period. In the year 2000, the US produced 1.4 million tonnes of peaches across 77,000 ha; in 2020, 560,000 tonnes of peaches were produced across 29,000 ha (FAOSTAT, 2020). This is in part due to declining profit margins, rising costs of labor and production costs for commercial peach growers across the US. An increased research and study of peach tree architecture and morphology will potentially reduce production costs, improve yields/land-area, and lead to the most economical use of resources.

Peach trees, like most fruit tree crops, require training systems. These training systems help to control and more efficiently direct the trees' innate growth, morphology and architecture as to increase fruit production, and regulate branching patterns and vigor. Establishing training systems however is laborious, time-consuming, and requires the trees to undergo regular pruning as additional maintenance. The above listed traits are becoming of greater importance from an agronomic perspective, due to the increasing costs of pruning and labor. This leaves ample room to investigate new methods of controlling and selecting for optimal architectural and morphological traits in peach trees. However, this process of adapting peach trees to training systems causes stress to tree and necessitates the need for regular pruning. Another issue that has been affecting US commercial growers in certain areas is a sub-optimal (or in some cases, negligible) increase in yields in the past 20-30 years (Marini and Sower, 2000). This nearly stagnant increase in peach yields for some growers has correlated to a similar stagnation of our orchard management and training systems. In the US, it seems that the majority of the training and orchard management systems we currently use, are the same ones since decades past, with little to no adaption or change occurring. These issues however may be addressed in unison with a further understanding of the various mechanisms controlling peach canopy architecture and morphology. Likewise, finding ways to categorize and quantify elements of tree architecture are vital to understanding such a highly quantitative and nuanced trait such as architecture and morphology. To do this however, first there needs to be a consistent and relatively user-friendly methodology to collect the required data.

In the most simplistic form, tree architecture can be described as the three-dimensional arrangement which is entirely comprised of the tree's above-ground surface: from the base of the stem to the top of the canopy (Lau et al., 2018). Tree architecture encapsulates several aspects of tree morphology and presents even more challenges to researchers who study it. While some might argue that the biggest hurdle is the highly quantitative nature of tree architecture, or the difficulty in accurately conceptualizing and modeling a tree's architecture, there is one larger dilemma: collecting the necessary data. This is due to the nature of the data, which is most commonly branching data. The current method of gathering branching data are manual, in-situ collection: measuring/counting tree branches and limbs via hand, usually with the assistance of several different people. This method of data collection is often laborious, time-consuming, and prone to human error (Bucksch, 2014). Our study investigates the use of terrestrial laser scanning (TLS) as a viable alternative approach. TLS is a type of remote sensing method that has been noted several times as an invaluable tool used to better quantify the architecture of tree canopies and assess their composition/biomass (Tanago et al., 2017). More specifically, terrestrial LiDAR (tLiDAR) scanning would be utilized to construct a point cloud representation of several peach trees, from which we could employ the use of several software modeling packages in order to collect the necessary branching data required, all in-silico. One such modeling software utilized in this study was *TreeQSM,* which was developed by distinguished researchers at Tampere University (Raumonen et al., 2013). *TreeQSM* allows us to create quantitative structural models (QSMs). These QSMs are constructed by fitting cylindrical overlays around the point cloud data provided from our peach tree tLiDAR scans. From these QSMs, we are then able to generate our in-silico biometric and branching data.

To summarize, the goal of our study is to investigate tLiDAR scanning, point cloud data analysis, and 3D modeling processes in order to create a feasible means of producing reliable in-silico branching data. We examined a number of different methods to validate our findings such as analyzing the residual ground truth data generated from our QSMs in comparison to the original point cloud data generated. We hope to corroborate our findings further by comparing in-situ field data and our in-silico generated data.

MATERIALS AND METHODS

3D scanning parameters and data collection

The in-silico and in-situ data for this comparison were collected at the Dempsey Farm at the University of Georgia in Griffin, GA. This site supports the Peach Research and Extension Orchard. The precise methods for this data collection were previously described in an earlier manuscript, but will be revisited here for clarification (Knapp-Wilson et al., 2022). Approx. 50 adult trees from the germplasm orchard were scanned while dormant in 2020, 2021 and 2022. These 50 adult peach trees represented 25 different cultivars (2 trees per cultivar) grafted onto 'Guardian' rootstock. Standard production and management guidelines were followed as recommended in the Southeastern Peach and Plum management guide (Blaauw et al., 2022). It is important to note that all the peach trees included in this study were planted in 2015 and managed using an open-vase training system. This results in roughly 3-5 'scaffolds' (large, 1^{st} order branches) being produced from the main trunk of the trees.

All the scans used in this study were taken with a FARO Focus3D X 330 laser scanner (Faro Technologies, Lake Mary, FL). Field scanning parameters were kept consistent between all years, with the scan duration being 11:28 per individual scan. The scans were 10240×4267 in relative dimensions, resulting in an approx. average of 43.7 million points being generated per scan (MPts). An approx. average of 0.56 cm was recorded as the distance between points in a 9.15 m cubic space (30 ft^3). Scans were taken between individual trees, on both sides of a respective row. This results in scans being collected approx. every 6.1 m. A minimum of 6 spherical targets were used every year during scanning. These targets helped with the registration process for cloud point reconstructions of our orchard. After a successful scan, three of the targets were moved forward while three remained behind as references. After successfully scanning a section of the orchard, the tLiDAR scans were saved and processed in SCENE v2020.0.1 (Faro Technologies, Lake Mary, FL). In SCENE, tree scans were processed using options to remove stray-points and dark-points within the scans. After finishing processing/registering the scans, point cloud data from individual trees were selected for reconstruction via *TreeQSM*. For this study, 4 adult trees from the first two rows of our germplasm collection were selected based on scan quality and unique architectural features. These trees, labeled as 1A, 6B, 11B, and 3A respectively, were collected from different peach cultivars. QSMs were generated for all 4 trees, alongside biometric and branching data in *TreeQSM*.

In-situ field branching data collection

In-situ data was also collected for adult germplasm trees in 2021 and 2022. This in-situ data consists of total number branching orders, as well as the number of branches per branching for each respective peach tree. Collection of in-situ data for both sets of trees occurred 1-2 weeks following tLiDAR scanning. The collection and counting of branches for the purpose of in-situ data collection followed procedures outlined by Lau et al. (2018). This was mostly done to be consistent with the standard practices of justifying bifurcations in branching, identification of branch nodes, continuation, and termination points of branching orders, and labeling of parent branches. The data collection was kept standard across all peach trees.

Optimization of 3D cylinder modeling pipeline *TreeQSM* for use in peaches

TreeQSM was developed by researchers in Finland for use in forestry. To this extent, it has been shown to work efficiently and reliably. However, adaptation of *TreeQSM* for use in horticulture; specifically in fruit trees, has seen little exposure. There have been previous studies done been modeling apple trees, although it appears our research is the first application of utilizing *TreeQSM* in modeling peaches (Zhang et al., 2020). The feasibility of using *TreeQSM* in order to accurately collect in-silico data for peach trees requires a more precise optimization of parameters than is what regularly used for forestry applications. These methods of parametrization will be followed as previously described in Knapp-Wilson et al. (2022). *TreeQSM* possesses several methods if directly inputting or changing parameters that are inherent to the modeling reconstruction process. In total there are five main parameters which effect different aspects of the cylindrical model overlay; however, three of these parameters are of major concern: PatchDiam1, PatchDiam2Min and PatchDiam2Max. While these three parameters result in the most drastic changes to how the model is created, all five of the parameters were ultimately optimized for use in peaches; specifically, in peach trees trained in an open-vase training system. Once the models are deemed within adequate parameters, 40 iterations of the modeling processes were executed. This produced 40 respective QSMs for each individual tree, and branching data values being collected per QSM per tree. Models created from this parameter optimization process are then tested and evaluated for accuracy to the original point clouds from which they are derived, with examples of both being shown in Figure 1. This is in order verify the model recreations are as perceivably true in their depiction of the original point cloud data as possible, which in turn validates the in-silico data when viewed in comparison to our field collected in-situ data.

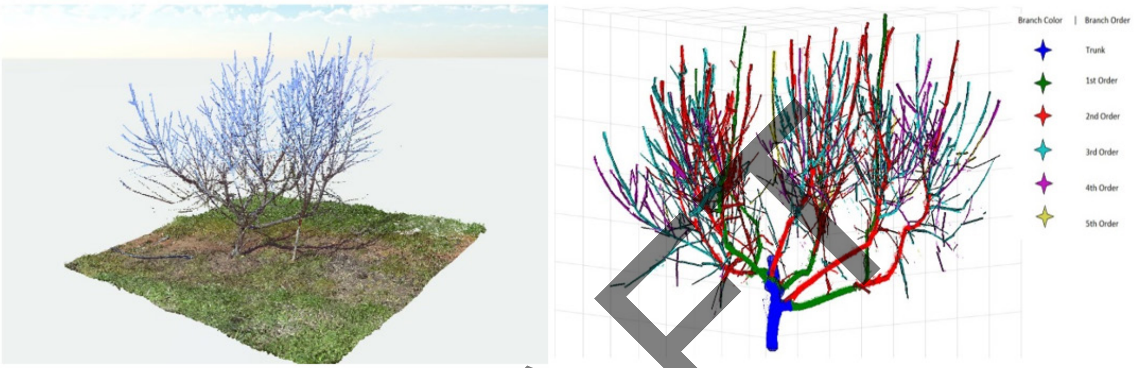

Figure 1. Modeling representations of adult germplasm tree 1A. Left is a basic point cloud reconstruction that was done in FARO software SCENE. Right is a quantitative structural model (QSM) created in *TreeQSM*.

Analysis of model-fitting metrics

To do this, we analyzed the residual ground truth data from our scans, as *TreeQSM* gives us a number of metrics to examine. However, the two metrics with the most significance to our study are: average distance (mm) from the original point cloud to the cylindrical model fitting, and the average surface coverage (%) of the cylindrical model in respects to the origin point cloud. Both of these metrics, avg. cylinder-point distance and avg. surface coverage are calculated for specific segments of every QSM generated for respective tree. *Treeqsm* also partitions the results from these metrics into variable section that correspond to different branching orders of the tree in addition to its entirety. These are sections are 'trunk', 'branch', '1branch, and '2branch'; which respectively correspond to measurements from trunk of the tree, all branching orders of the tree (excluding the trunk), 1st order branches, and 2nd order branches. This will be discussed further in the results section in more detail when going over Figure 2.

RESULTS AND DISCUSSION

QSM residual ground truth data

As mentioned previously we focused mainly on the two metrics from our residual data: avg cylinder-point distance and avg. surface coverage, show in Figure 2 below. The results in measurements for both metrics being created for the trunk, worth noting the parameters used when creating these QSMs, and the respectively generated residual data, were standard across all four-adult tress, across all 40x modeling iteration per tree. An example of residual data analysis is shown in Figure 2 for adult tree 1A. The average scores for all metrics are shown across the 40× iterations, with noticeable outliers (one shown in Figure 2), being present in both the findings of avg. surface coverage and avg. cylinder-point distance. Our findings showed that the average surface coverage percentages for both the trunk and 1st order segments of tree 1A were well above 70% (approx. 78 and 82%, respectively). These findings are above expected results for cylindrical model reconstructions. This is also in line with the lowest percent coverage segment 'branch' (the average of all higher order branches) being found to be approx. 50%, as it is the area in which our models encounter the highest percentage of stochasticism and branch occlusion during our modeling processes. The average measurements across all segments (i.e., entire tree averages) were also computed and evaluated for all four adult trees. These measurements, shown in Figure 3, were consistent with earlier preliminary findings when evaluating tree 1A in Figure 2. The four studied trees were found to be above 60%. Average point-cylinder distance across all adult trees was also found to be less than 7.5 mm. These findings give confidence in comparing in-situ and in-silico findings, since having averages less than 7.5 mm, even when considering the higher order branches, are great modeling standards. This is also considering that *TreeQSM* will generate models even with surface coverage %'s that are <20%. These results demonstrate levels of fidelity in the modeling accuracy when compared to the original point cloud data.

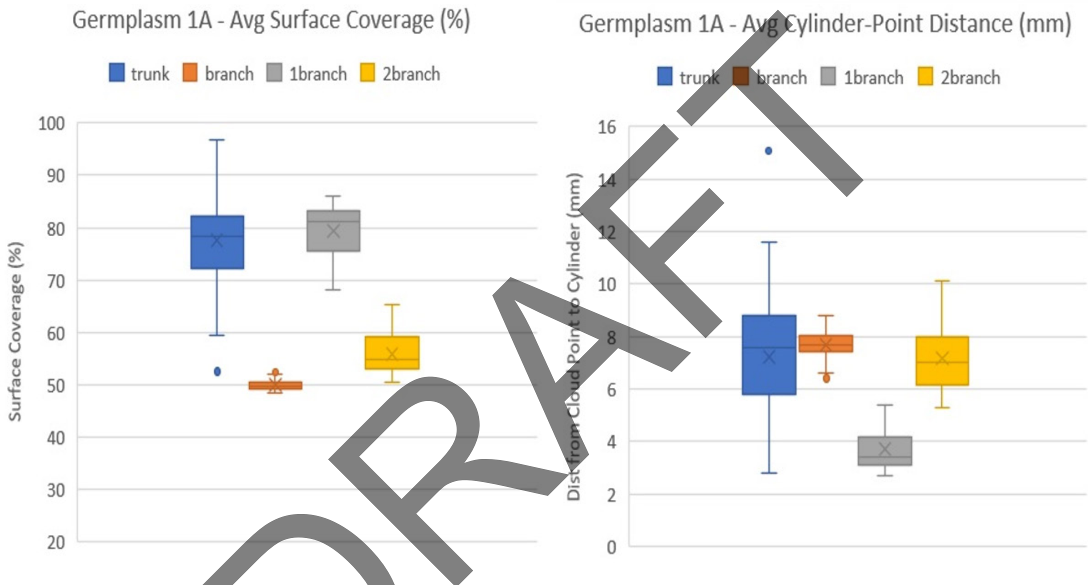

Figure 2. Box and whisker plot showing average cylindrical surface coverage (left) and average cylinder-point distance (right) measurements are shown for adult tree 1A. These scores were averaged across 40 modeling iterations, i.e., from 40 distinct QSMs. These metrics are used as a measure of ground truth error in order to investigate the accuracy of our models.

Next step: comparison of in-situ and in-silico branching data

After the analysis of QSMs created for the adult trees produced reliable results, our next and final step in verifying the in-silico data we collect for peach trees will be the comparison of in-situ and in-silico data. Preliminary analyses have already been done in younger peach trees, wherein no statistically significant difference was found between the in-silico and in-situ data collected for all respective young trees (data not shown). However, adult trees have considerably more branching density, higher branching orders, and overall more complexity in their morphologies when compared to younger trees. Therefore, comparisons between these two data sets (in-silico and in-situ) need to be considered over the range of multiple years, across multiple trees. This is our current step and focus of research. However, the preliminary findings in the young peach trees, as well as the results from our residual data analyses, give us confidence in results going forward.

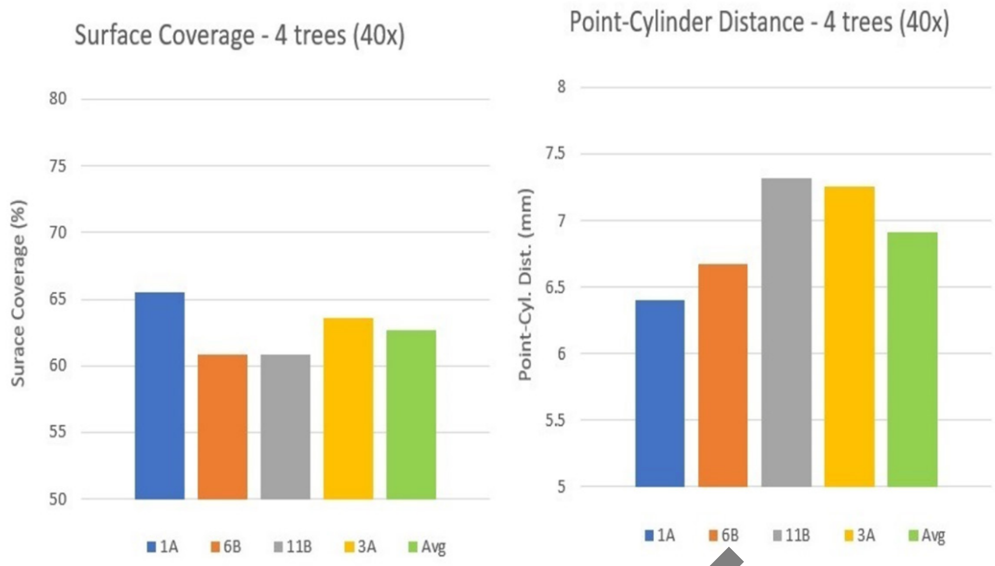

Figure 3. Bar graph showing average cylindrical surface coverage (left) and average cylinder-point distance (right) of entire trees, across all adult trees. Avg. surface coverage across all adult trees was above 60%, and all adult trees averaged under 7.5mm in point-cylinder distance. As in Figure 2, these scores were averaged across 40 modeling iterations, i.e., from 40 distinct QSMs.

CONCLUSIONS

To summarize, research into tree architecture (and by extension, tree morphology) can improve our understanding of these complex traits, which can lead to the development of better methodologies for shaping and controlling tree architecture than our current training systems. However, this research necessitates a change to the current method of branch data collection. Constructing an improved methodology which relies on reproducible and accurate in-silico data collection would greatly improve the ease of which branching data are collected. To this effect, the focus of our study was investigating the reliability of these models. This was done by comparing our models to the original point cloud data; effectively demonstrated by analyzing the residual data of our models over multiple iterations, across four adult trees. The results from these analyses showed reliable cylinder coverage percentages for all four trees. Similarly, the averaged point-cylinder distances for these adult trees were also found to be within credible levels (<7.5 mm). Overall, the results presented in this study support proceeding with comparisons of in-situ and in-silico data. This would be the final step in verifying the authenticity these quantitative structural models in terms of fidelity to their in-situ counterparts.

By demonstrating the fidelity of these models and the associated in-silico data, we can suggest the use of tLiDAR scanning and 3D reconstructive analyses from point cloud data as a practical alternative to collecting branching data manually. This would help in standardizing branching data collected, as well as making the process as a whole more reliable and freer from human error. The increased availability of branching data would in turn make researching tree architecture more accessible and reproducible; allowing for the sharing of point cloud data and 3D models from orchards across the world. Indeed, the availability of branching data are instrumental in decoding and identifying agronomically important architectural traits that can lead to increased resource efficiency, development of high-density orchards, agronomic production, and improved orchard management. It would seem appropriate that a method which allows the world to effectively collaborate in such research be critical for future advancements. In addition, combining all these qualities together opens new and alluring possibilities for future collaboration on agronomical production layouts and investigative genetic studies.

ACKNOWLEDGEMENTS

The research was supported by the NSF CAREER Award No. *1845760* and USDOE ARPA-E ROOTS Award Number DE-AR0000821 to A.B. as well as the Hatch Grant Initiative. Special thanks are given to the University of Georgia Institute of Plant Breeding, Genetics, and Genomics (IPBGG), as well as the members of both the Chavez Peach lab and the Bucksch Computational Plant Science lab. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the funders.

Literature cited

Blaauw, B., Brannen, P., Lockwood, D., Schnabel, G., and Ritchie, D. (2022). Southeastern Peach, Nectarine, and Plum Pest Management and Culture Guide (UGA Extension CAES) https://secure.caes.uga.edu/extension/publications/files/pdf/B%201171_16.PDF (accessed May 2, 2022).

Bucksch, A. (2014). A practical introduction to skeletons for the plant sciences. Appl. Plant Sci. *2* (*8*).

FAOSTAT. 2020. Statistics Division. Jan 5th 2021. http://www.fao.org/faostat/en/#home.

Knapp-Wilson, J., Bohn-Reckziegel, R., Bucksch, A., and Chavez, D.J. (2022). 3D phenotyping of peach tree canopy architecture using terrestrial laser scanning. NAPPN Conference, 22.-25. February 2022, Athens, GA.

Lau, A., Bentley, L.P., Martius, C., Shenkin, A., Bartholomeus, H., Raumonen, P., Malhi, Y., Jackson, T., and Herold, M. (2018). Quantifying branch architecture of tropical trees using terrestrial LiDAR and 3D modelling. Trees (Berl.) *32* (*5*), 1219–1231 https://doi.org/10.1007/s00468-018-1704-1.

Marini, R.P., and Sowers, D.S. (2000). Peach tree growth, yield, and profitability as influenced by tree form and tree density. HortScience HortSci *35* (*5*), 837–842 https://doi.org/10.21273/HORTSCI.35.5.837.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., and Lewis, P. (2013). Fast automatic precision tree models from terrestrial laser scanner data. Remote Sens. *5* (*2*), 491–520 https://doi.org/10.3390/rs5020491.

Tanago, J.G., Lau, A., and Bartholomeus, H. (2017). Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR. Methods Ecol. Evol.

Zhang, C., Yang, G., Jiang, Y., Xu, B., Li, X., Zhu, Y., Lei, L., Chen, R., Dong, Z., and Yang, H. (2020). Apple tree branch information extraction from terrestrial laser scanning and backpack-LiDAR. Remote Sens. *12* (*21*), 3592 https://doi.org/10.3390/rs12213592.