Estimation of Leaf Area Index using geospatial methods-A review

Oleh, T.C.[1] & Ajayi, O.G.[2]

[1.]Department of Surveying and Geoinformatics, Federal University of Technology, Minna,
[2.]Department of Land and Spatial Sciences, Namibia University of Sciences and Technology, Windhoek, Namibia
infotheosol@yahoo.com; gbenga.ajayi@futminna.edu.ng
correspondence email: infotheosol@yahoo.com

Abstract:

Leaf Area Index (LAI) is an essential vegetation leaf structure parameter for forest and agricultural ecosystems which can be estimated using different means. Remote sensing (RS) methods provide cost-effective alternatives to conventional field-based methods for LAI estimation. This paper presents the concept of LAI and methods for its estimation, with specific emphasis on geospatial methods. It also reviewed the concept of image classification in LAI estimation and the most commonly used image classification algorithms, which include random forest, supervised vector machine, artificial neural network, Bayesian, CARS-SPA, boosting, genetic, lookup tables, K-nearest neighbour, and modified K-nearest neighbour, while also highlighting some current research issues associated with LAI estimation from remotely sensed data. The findings of the study suggested that LAI can be accurately estimated from UAV imagery and that random forest, modified KNN, and conventional KNN algorithms are the most suitable classification algorithms for accurate LAI estimation using a remote sensing approach, especially when UAV images are used. The need for considering data science techniques for validation of the image classification approach for LAI estimation was recommended.

Keywords: Leaf Area Index, Classification, Algorithms, Geospatial Techniques, Forest Canopy.

INTRODUCTION

The scientific and efficient estimation of Leaf Area Index (LAI) is of significance for the evaluation of plant growth potential, as well as providing reliable technical support for the optimization of field management practices. The LAI used for determining forests, crops, climate, and soil is expressed as half the area of all leaves per unit area of the ground (Huan et al. 2022) LAI is a one-sided measure of leaf tissue per square inch of soil (Watson, 1947). A plant with a LAI of two (2) has the number of leaves that covers a given area twice. Different models use LAI to estimate fluxes between soil, canopy, and land-surface regions, and they do so in meters square (m2) (Bréada 2003). LAI can be estimated using either direct or indirect methods. The direct method of LAI estimation involves the physical measurement of leaf area within a particular area of interest which often requires destructive or allometric tree defoliation. This rigid procedure makes it difficult to apply the direct method or technique to large-scale studies and only suitable for the calibration of indirect methods (Bréada, 2003; Jonckheere et al. 2004). The direct method includes the clipping method which involves the collection of a sample of leaves from a plant canopy, and the measurement of their leaf area in a laboratory using a leaf area meter. The LAI is then calculated by dividing the total leaf area by the area of the ground sampled. Another method is the Point quadrat method which involves a device that is held at a fixed height above the canopy and has needles that randomly touch the leaves. The number of hits per unit area is recorded and used to estimate the LAI. The integrating sphere method is also a direct LAI estimation method that involves placing a plant inside a spherical chamber and measuring the light transmission through the chamber. The LAI is then calculated based on the amount of light that is absorbed by the leaves.

On the other hand, indirect method of LAI estimation involves the use of other parameters that are related to LAI, such as canopy reflectance, transmittance, or vegetation indices. These methods include: hemispherical photography, Light Detection and Ranging (LiDAR) which uses laser pulses to measure the distance between the sensor and the top of the canopy. The data can then be used to create a 3D image of the canopy, which can be used to estimate LAI, and spectral indices. The indirect method of LAI estimation is faster, easier, and automated, making it suitable for large geographic applications. The cost of indirect methods and the need for well trained personnel is the limitation of this method (Jonckheere et al. 2004).

Photogrammetry is one of the indirect, spatial based approaches to LAI estimation that uses a minimum of two (2) images of a spatial area to obtain three (3) dimensional (3D) information about the said area, which can be obtained by using information such as camera position, focal length, etc. Photogrammetry's primary purpose is to make use of two-dimensional images (2D) at different locations and re-produce into three-dimensional (3D) models of the same location. Recently, unmanned aerial vehicles (UAVs) have risen to prominence as a veritable photogrammetric method due to their cost-effectiveness in data collection and a high resolution compared to those from satellites and aircraft (Muhammad et al., 2020). The UAV photogrammetric process includes a remotely piloted aerial vehicle, a sensor payload, UAV flight planning, establishing and preparing ground control points (GCPs), check points (CPs), image processing with appropriate software, and map production and analysis as it relates to mapping projects (Muhammad et al. 2020). UAVs have proven to be reliable tool for the acquisition of spatial data and the extraction of geospatial data, such as LAI and image classification is one of the methods for extracting geospatial data from image-based spatial data.

Image classification assigns spectral classes to data classes. Spectral classes are groups of pixels that are similar in terms of their brightness values across the various spectral channels of data. An analyst can identify information classes in the image based on his knowledge and experience of the subject. For example, a remotely sensed image contains spectral signatures of many features that are present on the ground in terms of pixels of different values. When generating a thematic map, an interpreter or analyst sorts homogeneous groups of pixels with similar values and labels them as information classes such as water, agriculture, forest, etc. However, different spectral classes can be combined in one data class. Image classification consists of three main categories: supervised, unsupervised, and object-based image classification. Spectral signatures are unique, allowing one object to distinguish the other, which is the basis of classification in remote sensing, and they enable the extraction of LAI from remotely sensed data (Anupam and Krishna, 2012).

Some Image Classification Algorithms used for LAI Estimation

There are numerous image classification algorithms in use for the estimation of LAI but only a few of such algorithms for LAI estimation are discussed in this article. Table 1 presents the summary of the characteristics of the ten (10) selected image classification algorithms used predominantly for LAI estimation. The choice of these selected algorithms is informed by their frequency of use or implementation and their characteristics (Pirotti et al. 2016).

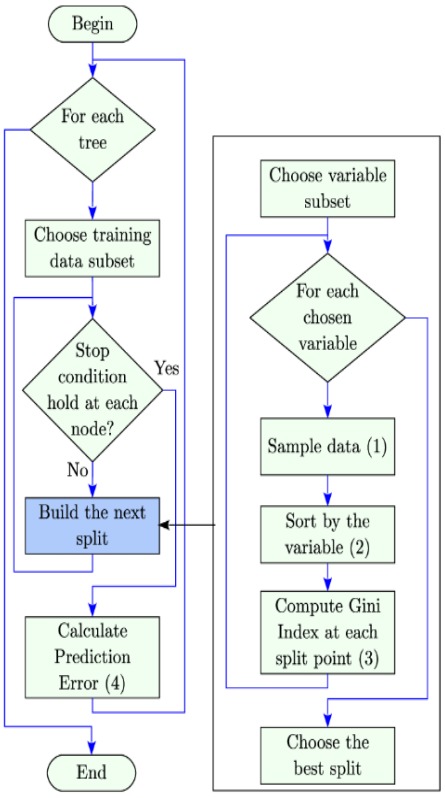
GEOSPATIAL METHODS OF LAI ESTIMATION

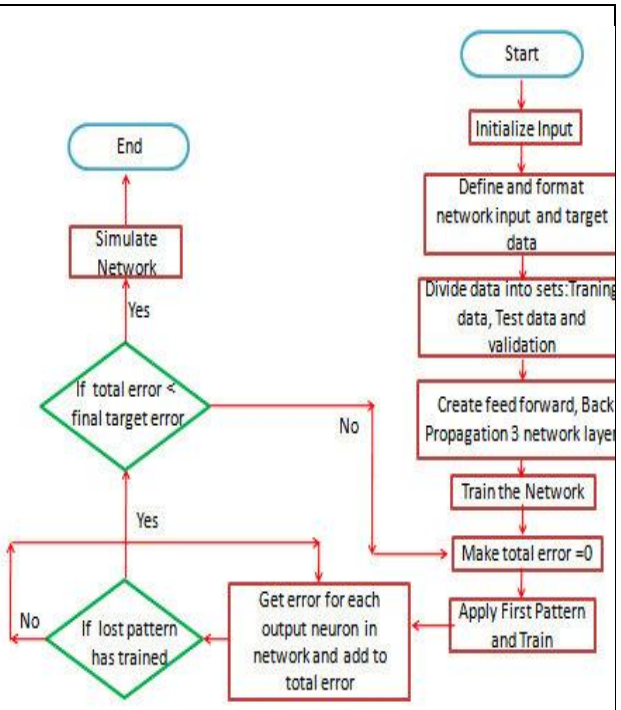
1. Direct measurement methods: It provides the most accurate LAI assessment and it is used for the evaluation of indirect remote sensing methods. It is laborious and less popular. Litterfall and biomass harvesting are examples of direct measurement (Beeck et al. 2019).

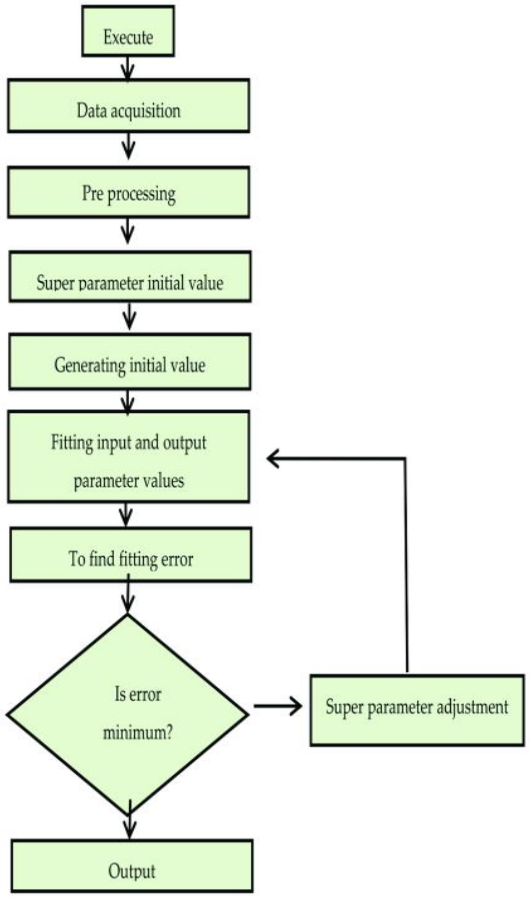
a. Litter fall: Deciduous plants lose their leaves annually; they carry same area as their annual leaf litter. Some leaves will fall mid-year before the canopy's maximum number of leaves forms It measures every area of leaf littered, this is one of the most accurate LAI methods. Meanwhile some of the littered leaves may blow away thereby reducing LAI measurement accuracy.

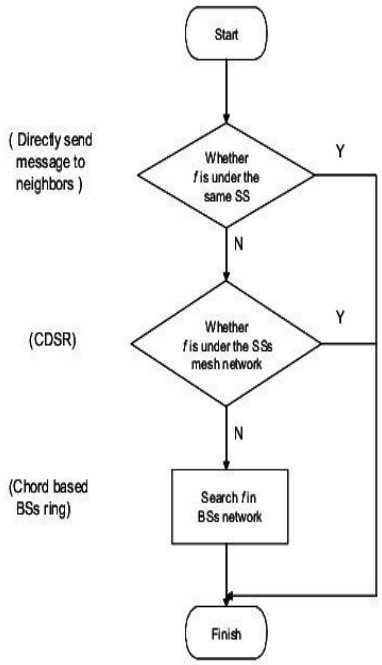
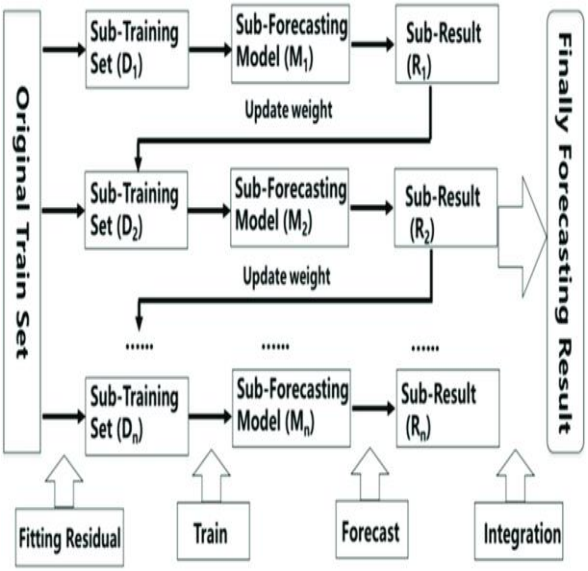
b. Biomass harvesting: It could measure leaf or needle area if LAI from a particular year was impossible. This is the most accurate but hardest way to estimate LAI for evergreen conifers. Stem area index calculates canopy woody element surface area (SAI). This destructive method affects no other plot measurements. This destructive method is the most accurate at estimating LAI, but it is rarely used unless it is the only option.

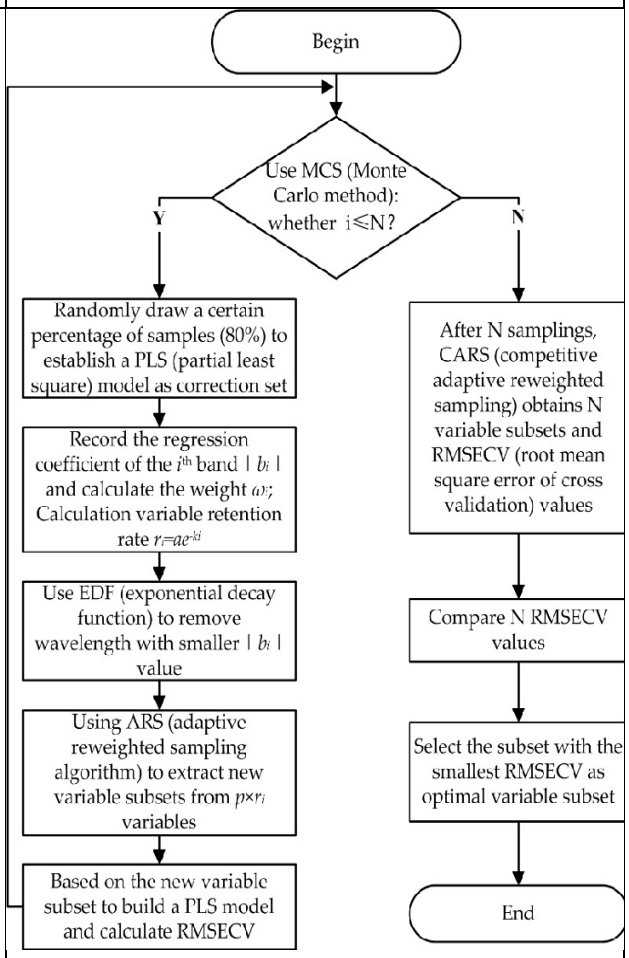
Table 1: Characteristics of selected image classification Algorithms.

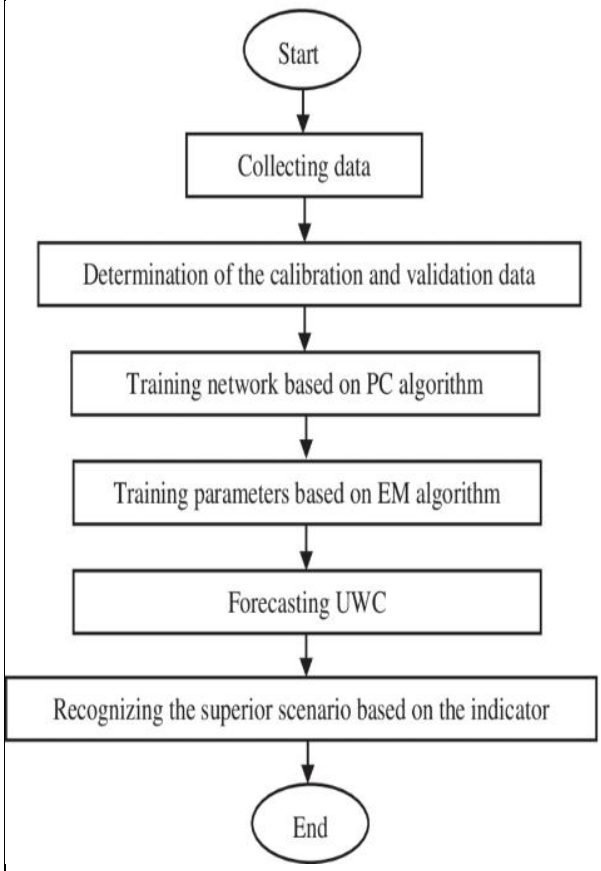
Image classification algorithms	References	Characteristics	Mathematical Formulation	Flow Chart Presentation of Process
Random Forest Algorithm (RFA).	Patel et al. (2019), Shuguo et al. (2020), Breiman, (2001).	-On large datasets, this set of facts is useful. - It can handle thousands of input variables without deleting anyone. - It estimates variables that are essential in image classification. -It generates a forests dataset that can be retrieved for future use (Breiman, 2001).	Given a training set X = x1, ..., xn with responses Y = y1, ..., yn, bagging repeatedly (B times) selects a random sample with replacement of the training set and fits trees to these samples: For b = 1, ..., B: Sample, with replacement, n training examples from X, Y; call these Xb, Yb. Train a classification or regression tree fb on Xb, Yb. After training, predictions for unseen samples x' can be made by averaging the predictions from all the individual regression trees on x': $\hat{f} = \frac{1}{B}\sum_{b=1}^{B} f_b(x')$ where $\hat{f}$ is the predictive estimation of the model.	 Source: Balli et al. (2019)

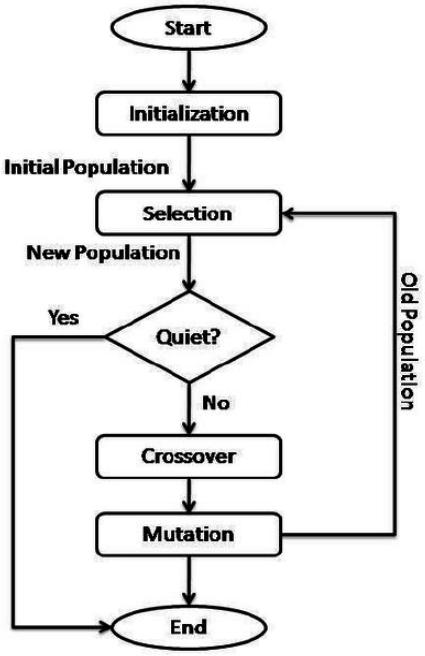
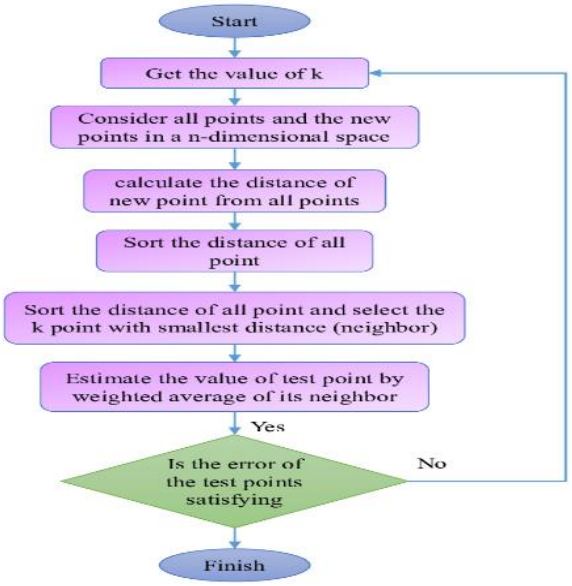
Artificial Neural Network (ANN).	Shunlin and Hongliang (2003), Zhang et al. (2011)	-For nonlinear systems dynamic findings and simulations, ANN has robust information modelling techniques that are used in a variety of ways. - Because of the ability to do approximation in a universal form and simple pattern, it allows the capture of complex nonlinear behaviour. - AN is also being used for research purposes (Nagy, 2007).	Building a neural net F so that ideally $F(x^{(i)}) = t^{(i)}$ where $x^{(i)}$ represent networks in neural net. However, typically we allow for error ϵ_i. Let $y^{(i)}$ denote the output of the neural net so that: $y^{(i)} = F(x^{(i)})$ and $t^{(i)} = y^{(i)} + \vec{\epsilon}_i$ We know that $y^{(i)}$depends on parameters, which are weights and biases, then it turns out as an optimization problem. We need to set up a neural net F that minimizes the error function, which is denoted below: $E = \frac{1}{N}\sum_{i=1}^{P}\|t^{(i)} - y^{(i)}\|^2$ where N is the number of training patterns and E is the error function.	 Source: Li et al. (2015).

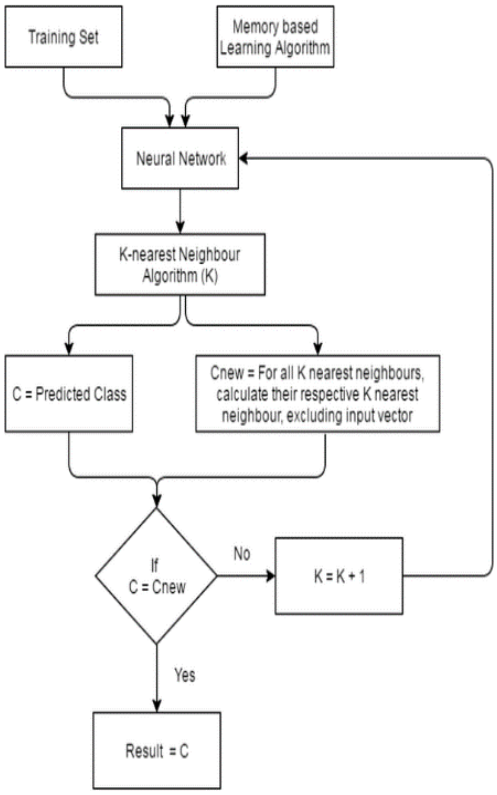
Support Vector Machines	Surya et al. (2006), Suthaharan and Suthaharan (2016).	-It uses hyperplanes to segment data that have been mapped to higher dimensions. - The information is mapped using a kernel. As it pertains to the data, various kernels are used. (Cortes and Vapnik, 1995).	Some mathematical kernel functions which you can use in SVM are given below: Polynomial kernel: $f(x_1, x_2) = (x_1^T . x_2 + 1)^d$. Here d is the degree of the polynomial, which we need to specify manually. Sigmoid kernel: We can use it as the proxy for neural networks. Equation is: $(x_1, x_2) = tanh(\propto x^T . y + x)$. It is just taking your input, and maps them to a value of 0 and 1 so that they can be separated by a simple straight line. RBF kernel is the most used kernel in SVM classifications, and it is depicted mathematically as follows: $f(x_1, x_2) = \frac{-\|x_1 - x_2\|^2}{2\sigma^2}$, $f(x_1, x_2) = e$. where, 1. 'σ' is the variance and our hyperparameter 2. $\|x_1 - x_2\|$ is the Euclidean Distance between two points x_1 and x_2. Another kernel is the Bessel function kernel: It is mainly used for eliminating the cross term in mathematical functions. Following is the formula of the Bessel function kernel: $k(x, y) = \frac{J_{v+1}(\sigma\|x-y\|)}{\|x-y\|^{-n(v+1)}}$. Finally, ANOVA Kernel performs well on multidimensional regression problems. The formula for this kernel function is: $k(x, y) = \sum_{k=1}^{n} exp\left(-\sigma(x^k - y^k)^2\right)^d$	 Source: Suthaharan and Suthaharan (2016)

Lookup Table Algorithm	Feng et al. (2006), Abebe (2006), Zhang et al. (1999)	LUTs are not used to creatively grade result, but rather they are used to make up a difference between a source and a result. (Friesen, 2011).	The mathematical representation provides a closed-form formula that allows us to compute the precise Lookup Table sizes for each Lookup Table transformation to minimize solution error as: $S_i = CS \Big/ \left(\sum_{j=1\ldots n} \sqrt{(M_j D_j / M_i D_i)} \right)$ $E_i = M_i D_i \left(\sum_{j=1\ldots n} \sqrt{(M_j D_j / M_i D_i)} \right) / CS$ Where: S_i: Estimation precise Lookup Table sizes E_i: Estimation solution error.	 Source: Gong et al. (2012).
Boosting	Shunlin et al. (2003), Huihui et al. (2019), Alzubi (2016).	-Boosting involves those algorithms that detects weaker learning classifiers when it comes to distribution and applying the same to the final stronger classifier. -When added to the final strong classifier, it is usually weighed in some ways, as it usually refers to weaker	(nction L: R2 → R, the gradient boosting algorithm iteratively constructs a model F:X→R to minimize the empirical risk ED[L(f(x),y)]. x, y ϵ X ϵ L , for L is a real number. At each iteration time (t), the model is updated as: $F^{(t)}(x) = F^{(t-1)}(x) + \varepsilon h^{(t)}(x)$. where $F^{(t)}(x)$ is estimated parameter of the model and $h^{(t)}(x)$ is a weak learner chosen from some family of functions.	

		learners' precision (Freund and Schapire, 1996).		Source: Alzubi (2016).
Competitive Adaptive Reweighted Sampling Projection Algorithm (CARS-SPA Algorithm)	Tang et al. (2014).	- It was first used in the selection of descriptive variables and SPA model it has minimal redundant data.	The use of CARS method begin with use of Monte Carlo method for sampling which can be represented mathematically given x- random variable at the random point in the interval [a,b], then $x_i = a + \sum_{k=1}^{n}(b-a)$. Secondly the Partial least squares method modelling is to decompose X and Y as follows $$X = TP^T + E$$ $$Y = UQ^T + F$$ In the formula, $T$and U are score matrices. P and Q are load matrices corresponding to X and Y respectively. E and F are residual matrixes. The relationship was established using Y =Tc +e = WXc + e = Xb + e. Where e is the prediction error, and $b = Wc = [b_1\ b_2 \ldots\ b_p]$is p-dimensional coefficient vector. $$r_i = ae^{-ki}$$ Where a and k are two constants. In the first sampling, all p variables are used for modeling, which means $r_1 = 1$. In the Nth sampling, only two variables are kept, so we get $r_N = {}^{2}/_{p}$. Under these two conditions, a and k can be calculated as: $a = ({}^{2}/_{p})^{{}^{1}/_{(N-1)}}$, k= $\frac{\ln({}^{2}/_{p})}{N-1}$, where ln represents the natural logarithm	 Source: Tang et al. (2014)

Bayesian Network Algorithm.	Terry et al. (2004), Jochem et al. (2015), liang et al. (2020).	- This is a joint probability distribution of a set of random variables with a potential mutual association. - It consists of nodes that represent random variables, edges between pairs of nodes that represent the nodes relationship and a conditional probability distribution within the nodes. - When observing new data, it models a posterior conditional probability distribution of the outcome variable(s) (Michal Horny, 2014).	The Bayesian Network Algorithm as the background from the framework of probabilities, you say that you have a prior distribution P(A), and next you wish to compute the posterior P(A | b). considering certain belief in the state of a particular variable, A. Next you receive the information that the state of the variable B is b, and you wish to use this information to update your belief in the state of A. For a Bayesian network exploiting the conditional independencies in causal networks, we get: Let CN be a causal network over U = {A1 , . . . , An}. Then P(U) has the following factorization: $$P(U) = \prod_{i=1}^{n} P\,[A_i \mid pa(A_i)]$$ The factorization above is the background for the concept Bayesian networks.	 Source: Sun et al. (2006)

Genetic Algorithm	Abraham et al. (2021), (Jenna, 2014)	- GA's are programming techniques that used to find the most efficient solution(s) to a given computational problem that maximizes or minimizes a particular function. - They are more effective than random search and complex search algorithms (Jenna, 2014).	Since genetic algorithms are designed to simulate a biological process, much of the relevant terminology is borrowed from biology. However, the entities that this terminology refers to in genetic algorithms are much simpler than their biological counterparts. The fitness function is the function that the algorithm is trying to optimize. If a problem has Npar -dimensions, then typically each chromosome is encoded as an Npar-element array chromosome = = [p1, p2, ..., pNpar]. The fitter a chromosome is, the more likely it is to be selected then the f is a non-negative fitness function, then the probability that a particular chromosome C is chosen to reproduce might be: $$P(C) = \left\| \frac{f(C)}{\sum_{i=1}^{N_{pop}} f(C_i)} \right\|$$	 Source: Lambora et al. (2019).
K-nearest neighbor (KNN) algorithm.	Mykola et al. (2020), (Samworth, 2012).	-It applies kernel functions in weighting neighbours with varying distances, so that not only kernel functions but also every monotonic decreasing function will work (Samworth, 2012).	The k-nearest neighbor classifier fundamentally relies on a distance metric. The better that metric reflects label similarity, the better the classified will be. It can be represented mathematically as: $$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$ Hence, function d(x,y) denote the classified estimation of metric at a distance point xi and yi.	 Source: Xiong and Yao (2021)

Modified K-nearest neighbour (MKNN) algorithm	Parvin et al. (2008), Mykola et al, (2020).	-It employs a kind of preprocessing on train data -It adds a new value called validity to train samples. -The validity takes into account the value of stability and robustness of any train sample regarding with its neighbors -Applying the weighted KNN employs validity as a multiplication factor yields a more robust classification.	The modified K-nearest neighbour (MKNN) can be represented mathematically $$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$ 2nd Step: Validity of Training Data Validity $$= \frac{1}{K}\sum_{i=1}^{K} S\left(|b|(x), |b|\left(N_{i(x)}\right)\right)$$ The S function is used to calculate the similarity between point x and the i-th data from the nearest, where a and b ∈ x. neighbor as follows: $$S(a,b) = \begin{cases} 1_{a=b} \\ 0_{a\neq b} \end{cases}$$ 3rd Step: Weight Voting $$W(x) = validalitas(x)x\frac{1}{d_{e+0.5}}$$	 Source: Sahu et al. (2018).

2. Indirect optical measurement methods: These methods measure light penetration through the canopy in the foliated stage and calculate the leaf area needed to produce the observed light above-under canopy relationship. Measurement systems struggle with penetration rates of about 5% with a leaf area index of six. Hemispherical photography, plant canopy analyzer, and sunscan ceptometer are some examples (Beeck et al. 2019).

a. Hemisherical photography: Hemispherical photography (fisheye or canopy photography) measures potential solar radiation and describes the plant canopy at 180 degrees. Since indirect optical methods estimate LAI, photographs have the highest-resolution optical data. The canopy images can measure canopy architecture metrics like gap fraction, sunfleck timing and duration, and others. They can be archived and re-analyzed as methods and software improve.

b. The plant canopy analyzer LAI-2000: Small hemispherical lenses detect light above and below the canopy. It uses the blue channel of the spectrum (320nm – 490nm), where the contrast between leaves and sky is greatest, and a second sensor to measure light above the canopy. This indirect optical method thrives in low light conditions and is not recommended for use in areas of very high sun light intensity, even though it is an efficient LAI estimation method. This method cannot measure LAI in humid conditions or when there is little or no sunlight.

C. The SunScan ceptometer: It uses high amount of direct light and is made possible by two sensors simultaneously above and below the canopy. It is one of many ceptometers available today, and it has been included to represent LAI measurements with ceptometers. The major strength of this method lies in its potential of measuring LAI accurately especially with the use of two (2) sensors above and below the canopy on a high intensity of sunlight. Thus, the main limitation is the inability to be used in poor atmospheric weather conditions like humid temperature, with little or no sunlight.

d. Airborne LiDAR: Airborne LiDAR is an indirect optical method. Knowing the time of light emission and the speed of light enhance the measurement outcome, backscattered data is used to determine the exact 3D position and other information about the reflecting material. Airborne LiDar is best for large areas from several square kilometres to entire regions, but it is expensive. LiDAR data provides a very accurate data for indirect LAI estimation (Bharat and Suddhasheel, 2017).

e. Unmanned Aerial Systems and satellite imageries: These involves gathering information about a place, object, or region without physical contact. This method accurately and cost-effectively calculates LAI by preparing their images, analysing and evaluating with algorithms. The main drawbacks of this method are low-resolution satellites, poor image analysis, and a lack of high-resolution sensors for UAV flights, especially in agriculture. Poor weather affects satellite and UAV reliability (Muhammad et al. 2020).

CONCLUDING REMARKS AND FUTURE OUTLOOK

Biophysical parameters like LAI predict stable agricultural growth and food availability in a long run. This review has provided concise information on LAI estimation with emphasis on its two basic approaches, while also presenting clear information on image classification algorithms and geospatial techniques that are used for estimating LAI. The review result shows that literature on LAI estimation concentrated on the indirect approach while validating the obtained results with direct methods. This direct method, which is invasive, still poses great threat to the healthiness and sustainability of plants because of the need to pluck off the leaves in the process of LAI estimation. Despite the rate of acceptance of data science techniques for validation of data experiments, little or no research has considered the approach for validation in LAI estimation. This method will not only provide accurate and reliable LAI estimation, but will ensure that the biophysical states of the leaves are not altered in the process of doing so. This will be the focus of our research in the near future.

REFERENCE

Abebe M. A. (2006). Retrieving Leaf Area Index of a Mature Norway Spruce Forest Stand From airborne Hyperspectral Image by inversion of the Dart Model. Thesis submitted in partial fulfillment of the degree of Master of Science in Geo-Information Science at Wageningen University and Research Centre, The Netherlands.

Abraham M., Francesca C., Giovanni C., Giacomo F., Eugenia C., Simonetta P., Mattia R., Emanuele S., Deodato T. and Claudia N. (2021). Biomass retrieval based on genetic algorithm feature selection and support vector regression in Alpine grassland using ground-based hyperspectral and Sentinel-1 SAR data. DOI: 10.1080/22797254.2021.1901063

Alzubi, J. A. (2016). Diversity-based boosting algorithm. International Journal of Advanced Computer Science and Applications 7(5):524 DOI:10.14569/IJACSA.2016.070570

Anupam A. and Krishna B. (2012). Processing and Classification of Remotely Sensed Images.

Balli, S., Sağbaş, E. A., and Peker, M. (2019). Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm. Measurement and Control - London- Institute of Measurement and Control- 52(3):002029401881369 DOI:10.1177/0020294018813692

Beeck M ,Carricondo c, Mauro f, , Roland M, Gielen B, Vitale D, Ceulemans R and Papale D.(2019). A comparison of different methods for assessing leaf area index in four canopy types DOI: 10.2478/forj-2019-0011, http://www.nlcsk.sk/fj/ Volume and Issue: Volume 65, Issue 2. Page range: 67 – 80

Bharat L. and Suddhasheel G. (2017). Airborne LiDAR Technology: A Review of Data Collection and ProcessingSystems. Proceedings of the National Academy of Sciences, India - Section A 87(4). DOI:10.1007/s40010-017-0435-9

Bréada, J. (2003). Ground-based measurements of leaf area index: a review of methods, instruments and current controversies. Journal of Experimental Botany, J Exp Bot. 2003 Nov;54(392):2403-17.doi: 10.1093/jxb/erg263.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Cortes C. and Vapnik V. (1995). Support-Vector Networks. AT and T Labs Holmdel, NJ 07733 USA. Machine Learning, 20, 273-297. http://dx.doi.org/10.1007/BF00994018

Feng D., Jing C., Stephen P., Mingzhen C. and Jan P. (2006) Algorithm for Global Leaf Area Index Retrieval Using Satellite Imagery. IEEE Transactions on Geoscience and Remote Sensing 44(8):2219 – 2229 DOI:10.1109/TGRS.2006.872100

Freund Y. and Schapire R. (2000). Game Theory, On-line Prediction and Boosting. Proceedings of the Nineth Annual Conference on Computational Learning Theory. DOI:10.1145/238061.238163

Friesen J. (2011). What is lookup table (LUT).

Gong, J., Liu, W., and Zhang, H. (2012). Multiple lookup table-based aes encryption algorithm implementation. Physics Procedia 25:842-847 DOI:10.1016/j.phpro.2012.03.166

Huihui M., Jihua M., Fujiang J., Qiankun Z and Huiting F (2019). Comparison of Machine Learning Regression Algorithms for Cotton Leaf Area Index Retrieval Using Sentinel-2 Spectral Bands. Appl. Sci. 2019, 9(7), 1459; https://doi.org/10.3390/app9071459

HuanY., Liping D., Xuan S., Ning W., CaixiaY., WeitongW. and Yun Z. (2022). Estimating leaf area index of maize using UAV-based digital imagery and machine learning methods. www.nature.com/scientificreports | https://doi.org/10.1038/s41598-022-20299-0

Jenna C. (2014). An Introduction to Genetic algorithms.

Jochem V., Gustau C., Jordi M., Juan P., Frank V., Jan G., and José M. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. ISPRS Journal of Photogrammetry and Remote Sensing DOI:10.1016/j.isprsjprs.2015.05.005

Jonckheerea I, Flecka S., Nackaertsa K., Muysa B., Coppina p., Weiss M., and Baretc F. (2004). Review of methods for in situ leaf area index determination Part I. Theories, sensors and hemispherical photography. Agricultural and Forest Meteorology 121(1):37-53 DOI:10.1016/j.agrformet.2003.08.001

Lambora A., Gupta, K., & Chopra, K. (2019). Genetic algorithm-A literature review. In 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon) DOI:10.1109/COMITCon.2019.8862255

Li, K., Hu, C., Liu, G., and Xue, W. (2015). Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis. Energy and Buildings, Energy and Buildings 108 DOI:10.1016/j.enbuild.2015.09.002

Michal H. (2014) Bayesian Networks. Technical Report No. 5.

Muhammad H., Anuar A. and Qudsia G. (2020). Impact of UAV Surveying Parameters on Mixed Urban Land-use Surface Modelling. International Journal of Geo-Information 9(11):656DOI:10.3390/ijgi9110656

Mykola K., Fugen J., Andrew R., Guangxing W., Hua L., and Hua S. (2020). A Modified KNN Method for Mapping the Leaf Area Index in Arid and Semi-Arid Areas of China pg 1-19. Remote Sensing 12(11):1884 DOI:10.3390/rs12111884

Nagy Z. (2007) Model based control of a yeast fermentation bioreactor using optimally designed artificial neural networks. Chemical engineering journal Vol127, No. 1-3, 95-109. journal ISSN :1385-8947DOI10.1016/j.cej.2006.10.015

Parvin H., Alizadeh H., and Bidgoli B. (2008). MKNN: Modified K-Nearest Neighbor. Proceedings of the World Congress on Engineering and Computer Science WCECS, San Francisco, USA.

Patel N., Subrata N., and Ritika S., (2019) Estimating leaf area index and light extinction coefficient using Random Forest regression algorithm in a tropical moist , India. DOI: 10.1016/j.ecoinf Ecological Informatics Volume 52, Pages 94-102

Pirotti F., Sunar F., and Piragnolo M (2016). Benchmark of machine learning methods for classification of a Sentinel-2 image. doi:10.5194/isprsarchives-XLI-B7-335-2016

Sahu, S. K., Kumar, P., and Singh, A. P. (2018). Modified K-NN algorithm for classification problems with improved accuracy. International Journal of Information Technology 10(3) DOI:10.1007/s41870-017-0058-z

Samworth R. J. (2012). Optimal Weight Nearest Neighbor Classifiers. University of Cambridge. The Annals of Statistics, Vol. 40, No. 5, 2733-2763 (31 pages) DOI:10.1214/12-AOS1049. https://www.jstor.org/stable/41806553

Sedat k., İlker E., Alkan G., and Muammer S. (2018). Artificial neural network models predicting the leaf area index: a case study in pure even-aged Crimean pine forests from Turkey. https://doi.org/10.1186/s40663-018-0149-8

Shuguo W., Liang L., Di G., Juan Y., Siyi Q., Liping D., Lu X., Lijuan W., Jianrong K. and Li L. (2020). Estimating Crop LAI Using Spectral Feature Extraction and the Hybrid Inversion Method. https://doi.org/10.3390/rs12213534

Shunlin L. and Hongliang F. (2003). Retrieving Leaf Area Index with a Neural Network Method: Simulation and Validation. 2052 IEEE Transactions on Geoscience and Remote Sensing 41(9):2052 – 2062 DOI:10.1109/TGRS.2003.813493

Sun, S., Zhang, C., and Yu, G. (2006). A Bayesian network approach to traffic flow forecasting. IEEE Transactions on intelligent transportation systems, 7(1), pp. 124-132. DOI: 10.1109/TITS.2006.869623

Surya S, Roger L and Nicolas H. (2006). Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. Remote Sensing of Environment 107 (2007) 348–361

Suthaharan, S., and Suthaharan, S. (2016). Support vector machine. Machine learning models and algorithms for big data classification: thinking with examples for effective learning, 56, p. 207-235. https://doi.org/10.1007/978-1-4899-7641-3_9

Tang, G., Huang, Y., Tian, K., Song, X., Yan, H., Hu, J., and Min, S. (2014). A new spectral variable selection pattern using competitive adaptive reweighted sampling combined with successive projections algorithm. Analyst, 139(19), pp. 4894-4902. DOI: 10.1039/c4an00837e

Terry C., Margret K., Sanchez-Azofeifa G., and Benoit R.(2004). Estimating Leaf Area Index From Satellite Imagery using Bayesian Networks. IEEE Transactions on Geoscience and Remote Sensing 43(8):1866 – 1873 DOI:10.1109/TGRS.2005.848412

Watson J. (1947). Leaf Area Index. Science Direct topics. Background introduction.

Xiong, L., and Yao, Y. (2021). Study on an adaptive thermal comfort model with K-nearest-neighbors (KNN) algorithm. Building and Environment, 202, pp. 108026. DOI 10.1016/j.buildenv.2021.108026

Zhang Y., Myneni R., Knyazikhin Y., Privette J., Running S., Nemani R., Tian Y., Wang Y., JMorisette J., Glassy J,. Lotsch A and Votava P. (1999). MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed By Vegetation (FPAR) Product (MOD 15). https://www.researchgate.net/publication/236770186

Zhang Z., Ma X., GuoShun L., Fang J., HongBo Q., Zhang Y., Shizhao L. and WenFeng S. (2011). A study on hyperspectral estimating models of tobacco Leaf Area Index.DOI: 10.5897/AJAR10.533. Advances in Food Science and Technology ISSN: 6732-4215 Vol. 8 (1), pp. 001-007, Available online at www.internationalscholarsjournals.org