



Article

A Complete Environmental Intelligence System for LiDAR-Based Vegetation Management in Power-Line Corridors

Domen Mongus ^{1,*} , Matej Brumen ¹, Danijel Žlaus ¹ , Štefan Kohek ¹ , Roman Tomažič ², Uroš Kerin ² and Simon Kolmanič ¹

¹ Faculty of Electrical Engineering and Computer Science, University of Maribor, Koroška Cesta 46, SI-2000 Maribor, Slovenia; matej.brumen@um.si (M.B.); danijel.zlaus@um.si (D.Ž.); stefan.kohek@um.si (Š.K.); simon.kolmanic@um.si (S.K.)

² ELES, d.o.o., Hajdrihova 2, SI-1000 Ljubljana, Slovenia; roman.tomazic@eles.si (R.T.); uros.kerin@eles.si (U.K.)

* Correspondence: domen.mongus@um.si; Tel.: +386-2-220-7408

Abstract: This paper presents the first complete approach to achieving environmental intelligence support in the management of vegetation within electrical power transmission corridors. Contrary to the related studies that focused on the mapping of power lines, together with encroaching vegetation risk assessment, we realised predictive analytics with vegetation growth simulation. This was achieved by following the JDL/DFIG data fusion model for complementary feature extraction from Light Detection and Ranging (LiDAR) derived data products and auxiliary thematic maps that feed an ensemble regression model. The results indicate that improved vegetation growth prediction accuracy is obtained by segmenting training samples according to their contextual similarities that relate to their ecological niches. Furthermore, efficient situation assessment was then performed using a rasterised parametrically defined funnel-shaped volumetric filter. In this way, $RMSE \approx 1$ m was measured when considering tree growth simulation, while a 0.37 m error was estimated in encroaching vegetation detection, demonstrating significant improvements over the field observations.

Keywords: LiDAR; vegetation management; digital twin; power-lines; encroaching vegetation detection; three growth simulation; environmental intelligence



Citation: Mongus, D.; Brumen, M.; Žlaus, D.; Kohek, Š.; Tomažič, R.; Kerin, U.; Kolmanič, S. A Complete Environmental Intelligence System for LiDAR-Based Vegetation Management in Power-Line Corridors. *Remote Sens.* **2022**, *13*, 5159. <https://doi.org/10.3390/rs13245159>

Academic Editor: Magaly Koch

Received: 22 September 2021

Accepted: 17 December 2021

Published: 20 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As electrification is becoming a pillar of social [1,2], economic [3] and environmental sustainability [2,4,5], power transmission lines are under increasing burden. While their performance monitoring and management have long been addressed through the concepts of so-called Smart Grids with the Internet-of-Things [6], benefits of environmental intelligence are still to be explored when addressing their physical safety [7]. As 30% of power outages are reportedly caused by weather conditions [8], 90% of which are attributed to tree-related incidents [9], vegetation management in power line corridors maintains a major challenge [7,8]. While it has already been shown that 6% improvement in reliability and 9% reduction in total costs is possible by only optimising tree trimming tasks [10], utilities still spend millions of dollars on vegetation management every year, making it one of the costliest activities in distribution asset management [7]. In addition, new challenges are now emerging related to mitigation of long-term negative impacts on biodiversity and ecosystems' sustainability [11,12]. Accordingly, digitalization in vegetation management has been explored increasingly in the last decade, within which the use of Light Detection and Ranging (LiDAR) has gained considerable attention [13].

Early researches into the subject were directed towards the extraction [14–18] and 3D reconstruction [19–23] of power lines from LiDAR data. Milzer and Briese (2004), for example, proposed minimum linkage clustering for pylons' extraction, while the extraction of power lines between them was achieved by using 2D Hough Transform, followed by 3D line fitting. On the other hand, McLaughlin [15] addressed the classification of transmission lines, vegetation and surfaces using eigenvalue decomposition. Hough Transforma-

tion and eigen-related features, together with surface-related, convex hull, echo-related, density-related and vertical profile-related features were explored by Kim et al. [16] for feeding a random forest classifier, while Cheng et al. [17] proposed Hough Transformation, eigenvalue and density features for clustering and 3D line fitting. Alternatively, Jwa and Sohn [20] proposed Hough Transformation for the definition of seed-points, while they achieved a reconstruction of power lines by using progressive line fitting combined with incremental segmentation. Alternatively, Guo et al. proposed Joint Boost-based classification of pylons [22] and power lines [21], and their reconstruction was achieved by using random sample consensus (RANSAC). Finally, Ortega et al. [23] performed a reconstruction of wires based on the catenary equation using particle swarm optimisation after an initial classification of pylons and wires, and their segmentation into individual conductors. Accordingly, while these traditional methods achieved mapping of pylons, followed by recognition of wires, more recent approaches focused on improving their performances [24] and extraction of more detailed information, such as, for example, the reconstruction of bundle conductors [25]. The latter was addressed by analysing single spans' fitting residuals, while each sub-conductor of bundle conductors was extracted by a projected dichotomy approach on the XOY and XOZ planes, and their reconstruction was performed by using a double-RANSAC with reported accuracies of above 90%. Contemporary methods can, thus, provide adequate support in power-line mapping tasks. However, assessment of power-line corridor clearance has, until recently, been a less frequently addressed research topic [13]. Despite the early recognised potentials [26], an efficient automatic LiDAR-based detection of clearance hazards (such as tree encroachment) has been reported only recently, with clearance measurement accuracy at the decimetre level [27], while over 95% accuracy of power-line and vegetation recognition for hazard detection was reported recently in [28]. The proposed process was based on dividing the large dataset into small manageable datasets for the generation of voxel-spaces, where separation of power lines from pylons and vegetation was achieved. Finally, the height and location of the extracted vegetation with reference to power lines were estimated for the detection of danger and clearance zones. Still, as argued in [13], fusion of multiple data sources can provide further benefits by reducing the monitoring costs, as well as improving temporal resolution (e.g., by fusion of aerial images [29]). In addition, data fusion can enhance monitoring with prediction capacities, as explored very recently with statistical predictions of tree-related power outages based on historical and weather data [30].

While predictive analytics is, thus, emerging as a new trend in effective power-line corridor management, improved vegetation growth simulations are needed, that are tuned to the exact ecological niches under inspection. Although many studies were conducted on the possible use of LiDAR data in vegetation and forest management, their primary focus remained on mapping the current state of the vegetation rather than using it for automated regression of vegetation growth. For example, Mongus and Žalik [31] proposed single tree-crown delineation and extraction of tree attributes by using Locally Fitted Surfaces (LoFS) and watershed segmentation, while Kolendo et al. [32] used a local maximum filter with growing region instead of LoFS. Donager et al. [33], on the other hand, proposed monitoring of forest structure by using eigenvalue and density metrics for the extraction of individual trees, with statistical analysis for assessing tree-level and stand-level attributes. Despite the fact that several studies have also indicated the possible use of LiDAR for predicting tree growth from multi-temporal datasets [34], such as by feeding the extracted information from LiDAR data into existing prediction engines [35], Random Forest or Linear Models [36], a complete data fusion stack that allows for vegetation growth simulations has not yet been introduced.

In this paper, we propose a new approach for achieving Environmental Intelligence in vegetation management using structured data fusion of LiDAR-generated data products with complementary thematic maps and administrative data sources, i.e., development of a digital twin [37]. Accordingly, following are the key scientific contributions of the paper:

- A complete LiDAR data processing pipeline for fusion of the derived data products (like digital terrain models, canopy height models and 3D data about power lines), with cadastral data and other important thematic maps for vegetation management, such as, for example, distribution of tree species and soil pH maps,
- An efficient approach for encroaching vegetation detection that enables accurate assessment of corridor clearance and provides future threat assessment, and
- A new data segmentation approach for learning vegetation growth simulation, with weak predictors tuned to specific ecological niches.

The rest of the paper is organised as follows: A new methodology for vegetation management is proposed in Section 2. Its results are presented in Section 3, while Section 4 concludes the paper.

2. Materials and Methods

2.1. Study Area and Data Source Preprocessing

In order to account for various testing conditions, an 18 km long corridor of the Slovenian national power transmission grid was selected. The corridor extends from the city of Nova Gorica to the town of Avče, and, thus, spans from the Sub-Mediterranean to the Alpine climate, and, accordingly, contains diverse forest stands with highly versatile terrain configurations. The terrain is also characterised by different soil qualities and soil pH levels, as well as sunlight conditions. In total, the power line corridor contains 104 power cables with a total span of approximately 168 km.

For the purposes of this study, two LiDAR data acquisitions were conducted in the years 2014 and 2018. As shown in Figure 1, the whole study area was divided into 1×1 km tiles, covering the total area of 24 km², while the protected area of the power-line corridor covered 2.72 km². In addition to LiDAR data, the auxiliary data sources used in this study are reported in Table 1.



Figure 1. Study area.

Table 1. List of auxiliary data sources.

Data-Source Description	Type	Year(s) of Acquisition	SPATIAL Resolution
Power transmission line axes	Geometry	2021	1.5 m
Mean amount of precipitation	Geometry	1981–2010	100 m
Mean air temperature	Geometry	1981–2010	100 m
Sunshine duration in summer	Geometry	1981–2010	100 m
Tree species distribution map	Raster	2020	10 m
Soil quality index	Raster	2011–2012	250 m
Soil pH	Raster	2011–2012	250 m

During source preprocessing, key LiDAR data products were generated, and their spatio-temporal data alignment with auxiliary thematic maps was achieved. For this purpose, a digital terrain model (DTM) and canopy height models (CHM) were generated from LiDAR data by using ground point and vegetation point classifications, as proposed in [38,39]. Note, however, that visual inspection and user refinements were necessary here, in order to correct inevitable inaccuracies introduced by automatic LiDAR data classification algorithms manually. Sampling of low ground points and high vegetation points into 1×1 km tiles was then performed with 0.5 m resolution, while data cleaning with interpolation of missing data and correction of tree heights was performed according to [40]. The obtained DTM was then subtracted from the digital surface model, as obtained from the sampled vegetation points, in order to define CHM. Nevertheless, due to the traditionally infrequent LiDAR data acquisitions, temporal alignment of the DTM and ground-truth CHM was achieved by considering the following additionally:

- Forest management activities conducted after LiDAR data were recorded;
- Vegetation growth up to the current date.

While the history of management activities was maintained with a log of the completed work orders and associated vector layers describing the region, date and type of cleaning tasks, tree growth predictions were used to estimate vegetation development. Temporal alignment was, thus, achieved iteratively (with 1 month temporal resolution), where, in each iteration, CHM and DTM were corrected in accordance with the power-line corridor management tasks from the previous month, followed by tree growth simulation, in order to approximate the current status of the vegetation. For this purpose, vegetation growth simulation was used, as proposed in [41]. Additionally, auxiliary raster data sources were resampled, according to [42,43], in order to achieve their alignment at 0.5 m resolution.

Finally, a detailed 3D geometry of the power transmission line was extracted from LiDAR data with 1.5 m resolution, as proposed by [24], resulting in a vector layer containing a little over a million 3D points. In order to provide a simplified assessment of their possible sagging, segments were also attributed with their distances from the transmission towers, as well as their voltage levels.

2.2. LiDAR Data Processing Framework for Vegetation Management

In this section, a complete data processing framework is presented for fusion of LiDAR derived data products (i.e., digital terrain and canopy height models), with auxiliary data sources (i.e., about forest species' distribution and environmental conditions), needed for ensuring accurate vegetation growth prediction and threat assessment in support of vegetation management in power-line corridors. For this purpose, we followed the JDL/DFIG (Joint Directors of Laboratories/Data Fusion Information Group) model, shown in Figure 2, which is considered to be a de-facto standard reference model for assessing features from heterogeneous data sources and streams. In addition to data preprocessing (i.e., Level 0), described in Section 2.1, it prescribes feature modelling over the following levels [44,45]:

- Level 1—Object assessment dealt with the definition of individual trees, their features, as well as the features of power-lines;
- Level 2—Situation assessment provided encroaching vegetation detection and risk assessment features;
- Level 3—Threat assessment integrated tree-growth predictions for the assessment of risk prognosis features;
- Level 4—Process refinement dealt with the management of other levels, recorded performance of the system, provided adaptive data acquisition and made decisions on how to improve the system efficiency;
- Level 5—User refinement dealt with knowledge management and visual analytics to support decision-making; while
- Level 6—Asset management, in our case, provided task scheduling by also considering available resources, legal constraints, and other operational factors.

As follows from the above, level 4 addressed overall system optimisations, while levels 5 and 6 were knowledge and vegetation management levels. Accordingly, we address in the continuation of this section levels 1 to 3, that provided environmental intelligence in support of these tasks.

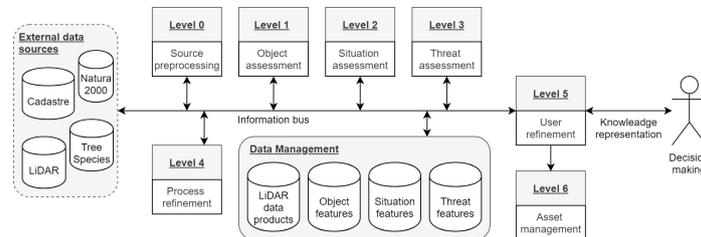


Figure 2. Overall concept of the data fusion framework.

2.2.1. Level 1—Object Assessment

During object assessment, we addressed the extraction of individual trees and fusion of auxiliary data sources about the environmental conditions relevant for their development. While the latter were used to provide contextual information for tree-growth prediction during the threat assessment, described in Section 2.2.3, spatial definition of individual trees was achieved with sufficient accuracy using single tree-crown delineation on a temporally aligned CHM. Single tree-crown assessment was then resolved by intersecting the resulting vector layer with auxiliary thematic maps, by which extraction of essential individual-tree parameters was brought about.

Formally, let a grid $G = \{p_i\}$ be given by a set of pixels $p_i = (x_i, y_i)$ that define the study area, while a general data layer is given by a mapping function $f : (x, y) \rightarrow \mathbb{R}$ acting on them, such as, for example, $CHM(p_i)$, and $DTM(p_i)$ defines the values of canopy height and digital terrain models at a given pixel p_i . A connected component $\mathcal{C}_t = \{p_i\} \subseteq G$ defines a delineation of a single tree-crown t , as obtained by [31]. Accordingly, an object assessment vector $\vec{A}(\mathcal{C}_t)$ that defines key environmental conditions associated with the tree \mathcal{C}_t is given by the following:

$$\vec{A}(\mathcal{C}_t) = \langle \widehat{CHM}(\mathcal{C}_t), \vec{P}(\mathcal{C}_t), SQ(\mathcal{C}_t), PH(\mathcal{C}_t), PR(\mathcal{C}_t), T(\mathcal{C}_t), DS(\mathcal{C}_t), LoFS(\mathcal{C}_t) \rangle, \quad (1)$$

where definitions of its components are provided in Table 2.

Table 2. Individual tree parameters, extracted from spatially aligned auxiliary thematic maps.

Name	Notation	Description	Data source
Tree height	\widehat{CHM}	The highest point within the extent of the tree crown.	CHM
Tree species	\vec{P}	Probabilities of the tree belonging to one of the three most common species in its extent	Tree species distribution map
Soil index	SQ	Average soil quality index within the extent of the tree crown	Soil quality index
Soil pH	PH	Average soil pH within the extent of the tree crown	Soil pH factors
Amount of precipitation	PR	The annual amount of precipitation per area of the tree crown	Mean amount of precipitation
Air temperature	AT	10 years average temperature within the extent of the tree crown	Mean air temperature
Sunshine duration	SD	10 years average sunshine duration in the area of the tree crown in summer	Sunshine duration in summer
Slope direction	$LoFS$	A slope normal, estimated by Locally Fitted Surface (LoFS) [38] to the area of the tree crown	Digital terrain model

2.2.2. Level 2—Situation Assessment

Following from the above, encroaching vegetation detection was performed using a 3D filter, defined by swept volume parametrisation [46]. In order to achieve its common definition regardless of power-line locations and attributes, a funnel-shaped volume generator was used, with parametrised width, height, and side-angle. While exact definitions of these may differ in accordance with the legislation, terrain configurations and other type-specifics of power-transmission lines, the following values were used in our case (see Figure 3a):

- The width of the filter was defined in accordance with the legislation, where 15 m was used for 110 kV transmission lines, while 40 m was used for higher voltage 210 kV and 400 kV power lines;
- The height of the filter was defined in accordance with the 3D shape of the lowest power-transmission line, ensuring at least 5 m clearance beneath it;
- The angle of the filter was fixed at 45° in order to prevent the risk of possible damage caused by falling high trees.

Accordingly, the volumetric filter definition was achieved by sweeping a generator along the power-transmission axes and storing it as a raster layer, which is given by a mapping function $\widetilde{CHM} : (x, y) \rightarrow \mathbb{R}$ that prescribes the maximum allowed vegetation height at a given pixel p_i . Note that \widetilde{CHM} was generated with 0.5 m resolution (see Figure 3b) in order to ensure its spatial alignment with the CHM . Thus, encroaching vegetation detection, together with the generation of a binary risk assessment function $RAM : (x, y) \rightarrow \{0, 1\}$, was achieved straightforwardly by pixel comparison. Formally, we have the following.

$$RAM(p_i) = \begin{cases} 1, & CHM(p_i) \geq \widetilde{CHM}(p_i) \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

where $p_i \in G$. Finally, vector layers were generated using isoline rendering [47], as shown in Figure 4.

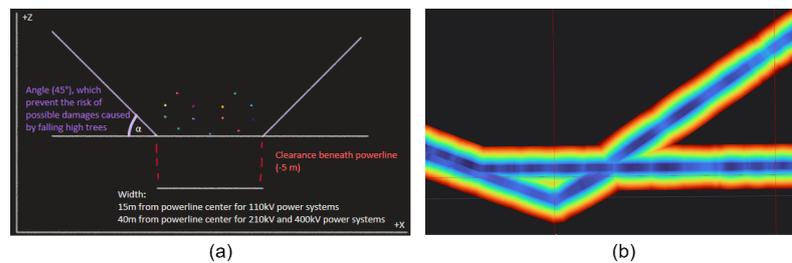


Figure 3. Filter definition (a) using a parameterized funnel-shaped generator, swept along the power-transmission line axis in order to obtain (b) a raster layer of the maximum allowed vegetation heights.

As, generally speaking, the growth of vegetation is observed at a much higher rate than structural changes of the power transmission lines, most of the computationally expensive tasks of the proposed approach were conducted during the preprocessing step. This concerns volumetric filter definition and estimation of \widetilde{CHM} , while only its pixel-comparison with CHM , as defined by Equation (2), was actually required during the processing. This proved to be useful, in particular when considering the predictions of encroaching vegetation on the simulated CHM and, thus, improving system performances during the threat assessment under different scenarios.

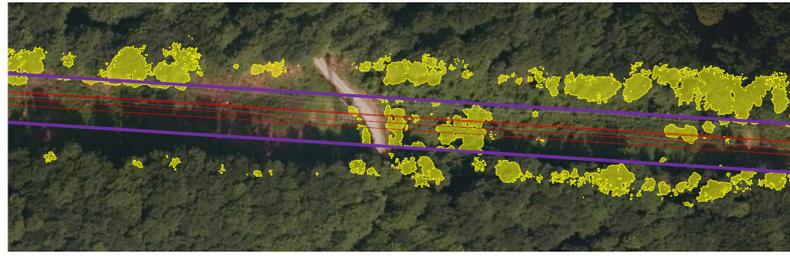


Figure 4. Encroaching vegetation detection.

2.2.3. Level 3—Threat Assessment

A new context-based ensemble regression for tree growth prediction is presented in this section that proved capable of dealing with various growing conditions and, together with the encroaching vegetation detection, allowed for holistic threat assessment. The rationale behind the approach was that trees growing within similar ecological niches behave similarly, and, thus, segmenting the learning data accordingly was expected to result in an improved prediction accuracy. Moreover, such an approach allowed for using fuzzy classification of individual trees' species (as provided traditionally by tree-species distribution maps), while also accounting for anisotropic tree-crown development (e.g., on forest-edges). This was achieved by considering parameters of individual trees from Table 2 as contextual features for segmentation of the learning data, while learning the regression model at the level of each individual CHM pixel (see Figure 5).

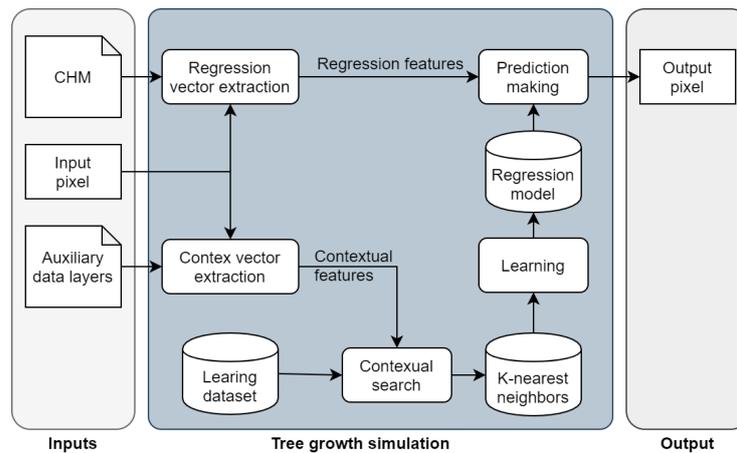


Figure 5. A flowchart of the proposed context-based ensemble regression.

Respectively, given a set of learning pixels $p_i = (x_i, y_i)$ with associated context vectors $C(p_i) = \vec{A}(\mathcal{C}_i)$, such that $p_i \in \mathcal{C}_i$, as defined by Equation (1), the context of a testing pixel p_t was defined by a subset $C^k[p_t] \subseteq C$ of k context vectors that are closest to $\vec{c}[p_t]$ according to some distance measurement. While an arbitrary mapping function $d : (\vec{c}[p_i], \vec{c}[p_t]) \rightarrow \mathbb{R}$ could be used for this purpose, contextual features were, in our case, of significantly different types and scales, and, thus, the $L1$ —norm was applied on ranked differences in feature-values rather than applying it on the feature-values themselves. Let a mapping function $rank_f : (\vec{c}(p_i)[f], \vec{c}(p_j)[f]) \rightarrow \mathbb{N}$ define a standard competition ranking of the difference between $\vec{c}(p_i)$ and $\vec{c}(p_j)$ according to the feature f , the used distance function d was defined formally as

$$d(\vec{c}(p_i), \vec{c}(p_j)) = \sum_f rank_f(\vec{c}(p_i)[f], \vec{c}(p_j)[f]). \quad (3)$$

Note that $d(\vec{c}(p_i), \vec{c}(p_j)) \neq d(\vec{c}(p_j), \vec{c}(p_i))$, while the regular difference between categorical (e.g., pH and soil quality levels), as well as numerical features (e.g., tree heights

and air temperature), was used for their ranking. On the contrary, ranking of vector-type features was achieved using angular distance.

Finally, regression model $R = \{R_{p_i}\}$ was defined by a set of weak regression models R_{p_i} , each associated with a testing pixel p_i . Two types of explanatory variables were used for this purpose, namely, a pixel and tree heights that related to the estimated increase in CHM due to the growth of the tree itself, and the heights of the neighbouring pixels that account for possible overgrowing of its surroundings. By defining the neighbourhood of $p_i \in C^k(p_i)$ using Cartesian product $W_S = \{x_i - S, \dots, x_i, \dots, x_i + S\} \times \{y_i - S, \dots, y_i, \dots, y_i + S\}$, where $S \geq 0$ specifies its size, a feature vector $\vec{v}(p_i)$ is given formally by an ordered set of CHM values as

$$\vec{v}(p_i) = \left(CHM[p_i], \max_{p_j \in W_1 \setminus W_0} CHM[p_j], \dots, \max_{p_j \in W_S \setminus W_{S-1}} CHM[p_j] \right), \tag{4}$$

where $W_S \setminus W_{S-1}$ refers to a set difference between two neighbourhoods W_S and W_{S-1} , as shown in Figure 6. Note, however, that by selecting the maximal CHM value within a given $W_S \setminus W_{S-1}$, orientation independent definition of weak regression models was achieved, while the terrain slope orientation and corresponding tree heights, together with fuzzy tree species classification and other contextual features, was already addressed during the learning data segmentation.

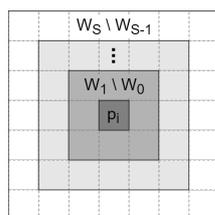


Figure 6. Definition of regression features.

In a sense, the developed prediction model can, thus, be considered as a KNN regression. This is an efficient lazy learning algorithm that, rather than providing a generalised model, uses all the training data for predicting the outcome of the target variable of testing samples. However, contrary to the traditional approach, the KNN search was achieved only on contextual features. Thus, although the implementation followed the optimisations proposed in [48], an arbitrary regression model could be applied for the actual predictions. As confirmed by the results (see Section 3), straightforward linear regression, applied on the $K = 100$ contextually most similar pixels, turned out to be the most efficient in our case.

3. Results

The proposed environmental intelligence system for vegetation management in power-line corridors (i.e., a digital twin) was implemented following high-performance monolithic architecture in the C++ programming language and deployed on three different test systems, specified in Table 3.

Table 3. Test systems.

Type	CPU	Number of Cores	Cache [MB]	Main Memory [GB]
Work-station	AMD® Ryzen™ Threadripper™ 1920X	12	39.1	64
Server	Intel® Xeon® E5-2650 v3	6	25	16
Laptop	Intel® Core® i7-9750HX	6	14	64

System validation was carried out on a test dataset P , containing $|P| \approx 100$ million pixels (as described in Section 2.1) from the following perspectives:

- Vegetation growth simulation accuracy was evaluated first, where Level 3 Threat assessment of data fusion was validated by pixel-comparison between the predicted CHM' and actual CHM using the root-mean-square error ($RMSE$) metric, defined as

$$RMSE = \sqrt{\frac{\sum_{p_i \in P} (CHM[p_i] - CHM'[p_i])^2}{|P|}}, \quad (5)$$

where $CHM'[p_i]$ was estimated by learning a weak regression model on $K = 100$ contextually closest pixels $P_{p_i}^{100} \subset (P \setminus \{p_i\})$ to a pixel p_i amongst all the pixels from the set $P \setminus \{p_i\}$;

- Encroaching vegetation detection validation was then achieved in order to validate data fusion Level 2 situation assessment by comparing the areas of detected risks with the history of the performed power-line corridor cleaning tasks; and
- System performances' assessment was finally carried out, where data preprocessing and object assessment, i.e., data fusion Levels 0 and 1, were evaluated additionally, and the overall data processing times are provided.

A detailed report of the obtained results is given in the continuation of this section.

3.1. Vegetation Growth Simulation Assessment

The validation of the vegetation growth simulation was carried out by comparing the accuracies and execution times achieved using three traditional regression approaches, namely, linear regression, KNN regression and artificial neural networks, with and without using learning data segmentation based on contextual features.

As follows from Table 4, notably higher execution times were measured when applying the proposed contextual segmentation of learning data, while this resulted in a decrease of $RMSE$ of all tested regression models, with an average improvement of over 10%. Figure 7 provides further details about the error distribution in comparison to the distribution of the measured vegetation growths. In all the cases, contextual segmentation managed to reduce error variance, as well as its range. However, as the majority of measured errors were within the $[-1, 1]$ range, while significantly larger $RMSE$ was measured, the presence of outliers was apparent. As linear regression with contextual segmentation of learning samples turned out to be the most accurate, showing little to no over-fitting, spatial distribution of errors obtained in this way is discussed further.

Table 4. Comparison of vegetation growth simulation accuracy and execution times on a test dataset, achieved by linear, KNN, and artificial neural network regressions, with and without using segmentation based on contextual features.

Regression Method	Segmentation	Execution Times [s]			RMSE
		Workstation	Server	Laptop	
Linear regression	No	170.3	293.5	188.3	1.16
KNN regression	No	179.6	306.4	198.6	1.38
Artificial neural network	No	342.7	586.7	379.0	1.36
Linear regression	Yes	602.6	1032.2	666.4	1.04
KNN regression	Yes	604.5	1035.2	668.1	1.29
Artificial neural network	Yes	774.2	1325.5	856.3	1.16
AVG	No	230.1	395.5	255.3	1.30
AVG	Yes	660.4	1130.1	730.3	1.16

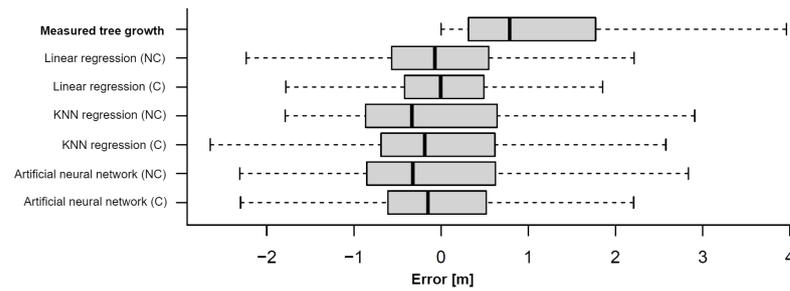


Figure 7. Distribution of measured errors achieved by the tested regression models with (C) and without (NC) using contextual segmentation of the learning data in comparison to measured tree growth.

A comparison, shown in Figure 8, shows that the simulated and ground-truth CHMs matched to a large extent (i.e., the yellow and orange colours in Figure 8d). However, a notable pattern of high error values is apparent, in particular, when considering forest edges. In cases of north and west edges, the proposed method overestimated the tree growth, while underestimations are more noticeable on the south and east edges. As similar, although less obvious, patterns can be noticed when considering the contours of dominant trees, larger errors were attributed to the predicted tree-crown expansion rather than to the predictions in vegetation growth.

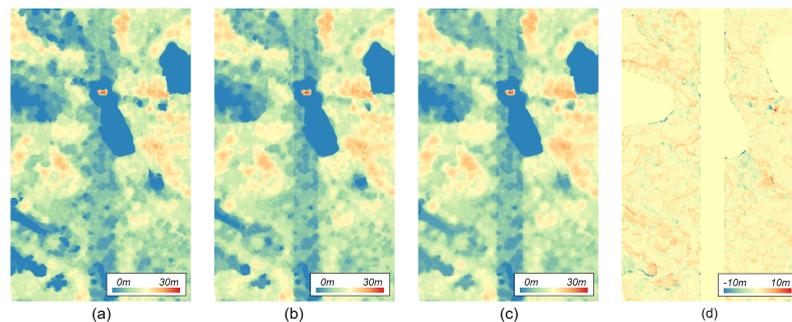


Figure 8. Spatial distribution of errors, where (a) shows an input CHM, (b) ground-truth, (c) simulated CHM and (d) the difference between the latter, where blue colours indicate low values, while red colours are used to display high values.

3.2. Encroaching Vegetation Detection

Encroaching vegetation risk assessment was conducted on a total of 168.5 km (as reported in Section 2.1) of power cables, modelled at 1.5 m resolution with exactly 1,000,325 points. The accuracy of the proposed approach was validated by comparing the detected risks with the field observations carried out by the asset management at the Slovenian national electricity transmission company (Eles d.o.o.). During the 2014–2017 period, 10 corridor clearances were carried out, covering a total area of approximately 1.9 km², with the largest covering the area of 7341 m² and the smallest related to an individual tree with the area of 22 m². Within these regions, the proposed method identified approximately 132 areas of encroaching vegetation with a total area of 0.5 km², with the area of individual regions ranging from 1415 m² to 0.25 m² (i.e., an individual pixel). As follows from the example shown in Figure 9, the reason for this lay in the fact that corridor clearances were carried out over the entire inner area of the power line corridor, not selectively on the detected encroaching vegetation.



Figure 9. Overlap between the areas of detected encroaching vegetation (red) and the clearance areas (green).

On the contrary, as indicated by the gray areas in Figure 10, the proposed method identified a number of threats to the safety of the power transmission lines outside of the clearance areas. In total, 449 such regions were detected, ranging in area from 228 m² to 0.25 m², with a total area of 1.8 km². Among these, 396 were smaller than 10 m² and can, thus, be attributed to the individual branches or their clusters rather than the actual trees. Providing these do not pose significant threats to the safety of the power transmission line, they can be thresholded straightforwardly during the post-processing. On the other hand, a large majority, namely 34 of the remaining 53 over-detected regions larger than 10 m², were detected on the forest edges, with individual branches posing a significant threat to the safety of the power transmission line, while the tree-tops themselves had not been detected as threatening. Their over-detection may, therefore, in a significant part, be attributed to the threat-assessment on the field, which is generally prone to errors. This was also confirmed by 16 over-detected trees behind the forest edge, which were, in the most part, not visible from the centre of the power line corridor. Finally, the remaining 3 over-detections were related to the misclassified LiDAR points.

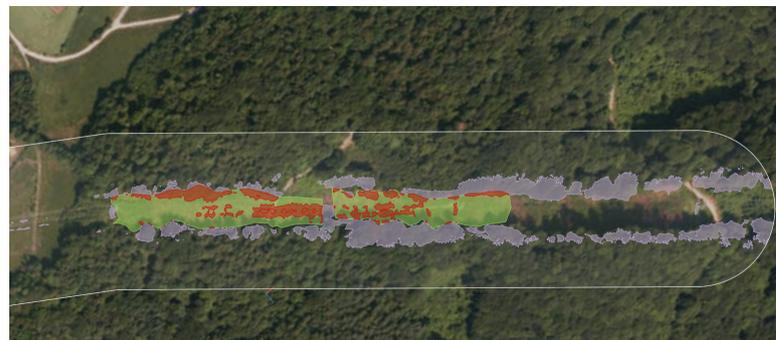


Figure 10. Over-detected encroaching vegetation (grey) outside of the clearance areas (green) with risk-areas (red).

Due to the reported uncertainty in the results, clearance measurement accuracy was additionally assessed on 5 individual trees, as described by [27]. Here, the average absolute accuracy of the detected encroaching vegetation equal to 19 cm was measured, with maximal error equal to 37 cm. While this is notably lower than reported in [27], where unmanned aerial vehicle LiDAR data acquisition was used, the LiDAR data were, in our case, recorded at a higher altitude (from a helicopter) and were, thus, of lower density.

In terms of data processing, on the other hand, the proposed approach consisted of a preprocessing step that included volumetric filter definition and its rasterisation, and the actual runtime processing with definition of situation assessment vector layer using ISO-lines. Accordingly, the experiments were conducted by running the whole encroaching vegetation risk assessment procedure 50 times, and measuring the system processing times during these steps.

Note that the reported time complexity of the proposed stages in Table 5 was derived by dividing the test power line into randomly defined segments and estimating the execution time per number of pixels used for encroaching vegetation detection. This procedure was repeated 50-times.

Table 5. Per-step time complexity analysis of encroaching vegetation detection per number of pixels N .

Step	Average Execution Time [s]			Time Complexity
	Workstation	Server	Laptop	
Volumetric filter definition	3.56	5.21	3.71	$O(N \log N)$
Filter rasterisation	2.51	7.23	2.48	$O(N)$
Encroaching vegetation detection	0.86	1.42	0.39	$O(N)$
Definition of ISO-lines	0.25	0.37	0.19	$O(N)$
Preprocessing (total)	6.07	12.44	6.19	$O(N \log N)$
Runtime (total)	1.11	1.79	0.58	$O(N)$
Total	7.18	14.23	6.77	$O(N \log N)$

3.3. System Performances

In addition to vegetation growth simulation and encroaching vegetation detection, the proposed approach considers source preprocessing and object assessment, conducted at data fusion Levels 0 and 1, accordingly. As these were achieved on the basis of previous studies, only their execution times are reported here, in order to provide a holistic validation of the proposed data fusion stack, while their accuracies, as reported in the original papers, were more or less confirmed during this study. The measured results, shown in Table 6, include the following:

- DTM generation, together with LiDAR ground point labelling, achieved during preprocessing as proposed by Mongus, Lukač, and Žalik in [38];
- CHM generation, including labelling of vegetation points, achieved during preprocessing as proposed by Horvat, Mongus, and Žalik in [39];
- Delineation of single tree-crowns, achieved during object assessment in accordance with the methodology proposed by Mongus and Žalik in [31];
- Calculation of slope direction, based on Locally Fitted Surfaces (LoFS), proposed by Mongus, Lukač, and Žalik in [38], achieved during the object assessment;
- Other processing steps, such as resampling of raster data used during preprocessing and estimations of intersections between different layers for extraction of contextual features during object assessment.

Table 6. Per-step time complexity analysis of data source preprocessing and object assessments' steps number of pixels N .

Step	Average Execution Time [s]		
	Workstation	Server	Laptop
DTM generation	2.79	5.21	3.08
CHM generation	6.45	10.77	7.22
Delineation of single tree-crowns	2.41	4.17	2.82
Calculation of slope direction	8.58	14.53	9.21
Other	1.02	1.78	1.16
Total	21.25	36.46	23.49

Together with vegetation growth simulations (see execution times in Table 4) and encroaching vegetation detection (see execution times in Table 5), the entire process of threat assessment on a 24 km² area was realised ≈ 10 min. On the workstation, that turned out to be the fastest, which was ≈ 20 min. On the server, that turned out to be the slowest amongst the tested computer systems. While the ranking of a test system was expected due to the computational power of their CPUs, the measured results clearly indicate that

the proposed approach can also be executed on less capable computer systems, such as, for example, data servers.

4. Discussion

As confirmed by the results, the proposed approach brings about an efficient environmental intelligence for improved vegetation management in power line corridors. By vegetation growth prediction and situation assessment, it enables predictive analytics to be achieved over the structured data fusion LiDAR derived data products with auxiliary thematic maps that followed the JDL/DFIG data fusion model. In this context:

- Spatio-temporal data alignment was achieved by data sub-sampling to a common resolution, while composing the current state CHM by adjusting it according to past clearance task and predicted vegetation growth from the time the LiDAR data were recorded. As previous studies have focused exclusively on mapping the state of power line corridors, the proposed approach offers improved monitoring capacities that prolong the relevance of the acquired data.
- Situation assessment based on parametric definition of a funnel-shaped volumetric filter can be achieved in preprocessing, which allows for fast encroaching vegetation detection. While the results achieved on higher high-altitude airborne LiDAR, showed slightly lower, yet comparable, accuracy to the related study performed on UAV acquired data, significant improvements in comparison to the field-based encroaching vegetation detection have been demonstrated.
- Threat assessment, enabled by vegetation growth prediction that utilises contextual segmentation of learning data for tuning weak regression models to specific ecological niches. While this improved prediction accuracy, the proposed approach provides the first attempt towards establishing a digital twin of the power line corridor ecosystem.

Despite the reported benefits of the proposed approach, the reported study provides only experimental validation, while additional test areas need to be included during system operation in order to achieve its demonstration in an operational environment. Moreover, despite $RMSE \approx 1$ m was achieved on all test systems, proving the reliability of the proposed data fusion stack, it still leaves room for improvement. Notably, significantly lower accuracy in the predictions of tree-crowns' expansions were measured in comparison to the predictions of tree growth, with spatial distribution of errors indicating its subjection to sunlight conditions. As typically overestimations occur on the north and west forest edges, underestimations were more characteristic on the south and east sides. While this can be compensated straightforwardly by an asymmetric filter definition that imposes stricter conditions on one side than the other, accordingly, the actual solution to this issues may require introduction of orientation depended regression features, or additional contextual features. As this requires an in-depth analysis of the impacts of tree-crown delineation and extraction of individual features on prediction accuracy by using possible feature learning, together with the assessment of the method's sensitivity to the parameter K , it is considered to be beyond the scope of this paper, and will be addressed in our future work. Furthermore, as the behaviour of each individual pixel was modelled with a dedicated weak prediction model, clustering of samples based on their contextual features may significantly speed-up the simulation's learning process. However, its impact on the accuracy should be studied. Finally, while the accuracy of encroaching vegetation detection is expected to improve with higher resolution datasets, an appropriate post-processing of the detected hazards is still required. In order to meet asset management requirements (i.e., data fusion Level 6), this should account for legal restrictions and cost requirements of power line corridor clearance tasks that will enable their optimal grouping, prioritising, and scheduling.

Author Contributions: Conceptualisation, D.M.; methodology, D.M.; software, M.B.; validation, R.T. and U.K.; formal analysis, D.Ž.; investigation, S.K. and Š.K.; resources, D.Ž.; data curation, R.T. and U.K.; writing—original draft preparation, D.M.; writing—review and editing, Š.K.; visualisation, D.M., Š.K. and M.B.; supervision, S.K.; project administration, D.M.; funding acquisition, R.T. and U.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Slovenian Research Agency under Grants P2-0041 and L7-2633 and ELES d.o.o.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Special thanks go to the co-funders and users of the developed technology, the Slovenian national power-network operator ELES d.o.o.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chaurey, A.; Ranganathan, M.; Mohanty, P. Electricity access for geographically disadvantaged rural communities —technology and policy insights. *Energy Policy* **2004**, *32*, 1693–1705. [[CrossRef](#)]
2. Garces, E.; Tomei, J.; Franco, C.J.; Dynner, I. Lessons from last mile electrification in Colombia: Examining the policy framework and outcomes for sustainability. *Energy Res. Soc. Sci.* **2021**, *79*, 102156.
3. Mandelli, S.; Barbieri, J.; Mereu, R.; Colombo, E. Off-grid systems for rural electrification in developing countries: Definitions, classification and a comprehensive literature review. *Renew. Sustain. Energy Rev.* **2016**, *58*, 1621–1646. [[CrossRef](#)]
4. Boliko, C.M.; Ialnazov, D.S. An assessment of rural electrification projects in Kenya using a sustainability framework. *Energy Policy* **2019**, *133*, 110928. [[CrossRef](#)]
5. Mukhtar, M.; Obiora, S.; Yimen, N.; Quixin, Z.; Bamisile, O.; Jidele, P.; Irivboje, Y.I. Effect of inadequate electrification on Nigeria's economic development and environmental sustainability. *Sustainability* **2021**, *13*, 2229. [[CrossRef](#)]
6. Li, Y.; Cheng, X.; Cao, Y.; Wang, D.; Yang, L. Smart choice for the smart grid: Narrowband Internet of Things (NB-IoT). *IEEE Internet Things J.* **2017**, *5*, 1505–1515. [[CrossRef](#)]
7. Dokic, T.; Kezunovic, M. Predictive risk management for dynamic tree trimming scheduling for distribution networks. *IEEE Trans. Smart Grid* **2018**, *10*, 4776–4785. [[CrossRef](#)]
8. Haes Alhelou, H.; Hamedani-Golshan, M.E.; Njenda, T.C.; Siano, P. A survey on power system blackout and cascading events: Research motivations and challenges. *Energies* **2019**, *12*, 682. [[CrossRef](#)]
9. Wanik, D.; Parent, J.; Anagnostou, E.; Hartman, B. Using vegetation management and LiDAR-derived tree height data to improve outage predictions for electric utilities. *Electr. Power Syst. Res.* **2017**, *146*, 236–245. [[CrossRef](#)]
10. Kuntz, P.A.; Christie, R.D.; Venkata, S.S. Optimal vegetation maintenance scheduling of overhead electric power distribution systems. *IEEE Trans. Power Deliv.* **2002**, *17*, 1164–1169. [[CrossRef](#)]
11. Eldegard, K.; Eyitayo, D.L.; Lie, M.H.; Moe, S.R. Can powerline clearings be managed to promote insect-pollinated plants and species associated with semi-natural grasslands? *Landsc. Urban Plan.* **2017**, *167*, 419–428. [[CrossRef](#)]
12. Çoban, S.; Balekoğlu, S.; Özalp, G. Change in plant species composition on powerline corridor: A case study. *Environ. Monit. Assess.* **2019**, *191*, 200. [[CrossRef](#)] [[PubMed](#)]
13. Matikainen, L.; Lehtomäki, M.; Ahokas, E.; Hyypä, J.; Karjalainen, M.; Jaakkola, A.; Kukko, A.; Heinonen, T. Remote sensing methods for power line corridor surveys. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 10–31. [[CrossRef](#)]
14. Melzer, T.; Briese, C. Extraction and modeling of power lines from ALS point clouds. In Proceedings of 28th Workshop of the Austrian Association for Pattern Recognition, Hagenberg, Austria, 22–23 May 2004; pp. 47–54.
15. McLaughlin, R.A. Extracting transmission lines from airborne LIDAR data. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 222–226. [[CrossRef](#)]
16. Kim, H.B.; Sohn, G. Point-based classification of power line corridor scene using random forests. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 821–833. [[CrossRef](#)]
17. Cheng, L.; Tong, L.; Wang, Y.; Li, M. Extraction of urban power lines from vehicle-borne LiDAR data. *Remote Sens.* **2014**, *6*, 3302–3320. [[CrossRef](#)]
18. Zhu, L.; Hyypä, J. Fully-automated power line extraction from airborne laser scanning point clouds in forest areas. *Remote Sens.* **2014**, *6*, 11267–11282. [[CrossRef](#)]
19. Jwa, Y.; Sohn, G.; Kim, H. Automatic 3d powerline reconstruction using airborne lidar data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2009**, *38*, W8.
20. Jwa, Y.; Sohn, G. A piecewise catenary curve model growing for 3D power line reconstruction. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1227–1240. [[CrossRef](#)]
21. Guo, B.; Li, Q.; Huang, X.; Wang, C. An improved method for power-line reconstruction from point cloud data. *Remote Sens.* **2016**, *8*, 36. [[CrossRef](#)]
22. Guo, B.; Huang, X.; Li, Q.; Zhang, F.; Zhu, J.; Wang, C. A Stochastic Geometry Method for Pylon Reconstruction from Airborne LiDAR Data. *Remote Sens.* **2016**, *8*, 243. [[CrossRef](#)]
23. Ortega, S.; Trujillo, A.; Santana, J.M.; Suárez, J.P.; Santana, J. Characterization and modeling of power line corridor elements from LiDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 24–33. [[CrossRef](#)]

24. Awrangjeb, M. Extraction of power line pylons and wires using airborne lidar data at different height levels. *Remote Sens.* **2019**, *11*, 1798. [[CrossRef](#)]
25. Zhou, R.; Jiang, W.; Jiang, S. A novel method for high-voltage bundle conductor reconstruction from airborne LiDAR data. *Remote Sens.* **2018**, *10*, 2051. [[CrossRef](#)]
26. Ko, C.; Rimmel, T.K.; Sohn, G. Mapping tree genera using discrete LiDAR and geometric tree metrics. *Bosque* **2012**, *33*, 313–319. [[CrossRef](#)]
27. Chen, C.; Yang, B.; Song, S.; Peng, X.; Huang, R. Automatic clearance anomaly detection for transmission line corridors utilizing UAV-Borne LIDAR data. *Remote Sens.* **2018**, *10*, 613. [[CrossRef](#)]
28. Munir, N.; Awrangjeb, M.; Stantic, B. An improved method for pylon extraction and vegetation encroachment analysis in high voltage transmission lines using LiDAR data. In Proceedings of the 2020 Digital Image Computing: Techniques and Applications (DICTA), Melbourne, Australia, 29 November–2 December 2020.
29. Mills, S.J.; Castro, M.P.G.; Li, Z.; Cai, J.; Hayward, R.; Mejias, L.; Walker, R.A. Evaluation of aerial remote sensing techniques for vegetation management in power-line corridors. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3379–3390. [[CrossRef](#)]
30. Hartling, S.; Sagan, V.; Maimaitijiang, M.; Dannevik, W.; Pasken, R. Estimating tree-related power outages for regional utility network using airborne LiDAR data and spatial statistics. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *100*, 102330. [[CrossRef](#)]
31. Mongus, D.; Žalik, B. An efficient approach to 3D single tree-crown delineation in LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 219–233. [[CrossRef](#)]
32. Kolendo, L.; Kozniowski, M.; Ksepko, M.; Chmur, S.; Neroj, B. Parameterization of the Individual Tree Detection Method Using Large Dataset from Ground Sample Plots and Airborne Laser Scanning for Stands Inventory in Coniferous Forest. *Remote Sens.* **2021**, *13*, 2753. [[CrossRef](#)]
33. Donager, J.J.; Sánchez Meador, A.J.; Blackburn, R.C. Adjudicating Perspectives on Forest Structure: How Do Airborne, Terrestrial, and Mobile Lidar-Derived Estimates Compare? *Remote Sens.* **2021**, *13*, 2297. [[CrossRef](#)]
34. Hopkinson, C.; Chasmer, L.; Hall, R. The uncertainty in conifer plantation growth prediction from multi-temporal lidar datasets. *Remote Sens. Environ.* **2008**, *112*, 1168–1180. [[CrossRef](#)]
35. Hudak, A.T.; Evans, J.S.; Stuart Smith, A.M. LiDAR utility for natural resource managers. *Remote Sens.* **2009**, *1*, 934–951. [[CrossRef](#)]
36. Zhao, K.; Suarez, J.C.; Garcia, M.; Hu, T.; Wang, C.; Londo, A. Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux. *Remote Sens. Environ.* **2018**, *204*, 883–897. [[CrossRef](#)]
37. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2021**, *58*, 346–361. [[CrossRef](#)]
38. Mongus, D.; Lukač, N.; Žalik, B. Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 145–156. [[CrossRef](#)]
39. Horvat, D.; Žalik, B.; Mongus, D. Context-dependent detection of non-linearly distributed points for vegetation classification in airborne LiDAR. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 1–14. [[CrossRef](#)]
40. Khosravipour, A.; Skidmore, A.K.; Isenburg, M.; Wang, T.; Hussin, Y.A. Generating pit-free canopy height models from airborne lidar. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 863–872. [[CrossRef](#)]
41. Strnad, D.; Kohek, Š.; Kolmanič, S. Fuzzy modelling of growth potential in forest development simulation. *Ecol. Inform.* **2018**, *48*, 80–88. [[CrossRef](#)]
42. Rees, W. The accuracy of digital elevation models interpolated to higher resolutions. *Int. J. Remote Sens.* **2000**, *21*, 7–20. [[CrossRef](#)]
43. Li, Z. Multi-scale modelling and representation of geospatial data. In *Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences: 2008 ISPRS Congress Book*; CRC Press: Boca Raton, FL, USA, 2008; pp. 283–296.
44. Meng, T.; Jing, X.; Yan, Z.; Pedrycz, W. A survey on machine learning for data fusion. *Inf. Fusion* **2020**, *57*, 115–129. [[CrossRef](#)]
45. Munir, A.; Blasch, E.; Kwon, J.; Kong, J.; Aved, A. Artificial Intelligence and Data Fusion at the Edge. *IEEE Aerosp. Electron. Syst. Mag.* **2021**, *36*, 62–78. [[CrossRef](#)]
46. Aigner, M.; Heinrich, C.; Jüttler, B.; Pilgerstorfer, E.; Simeon, B.; Vuong, A.V. Swept volume parameterization for isogeometric analysis. In *IMA International Conference on Mathematics of Surfaces*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 19–44.
47. Bajaj, C.L.; Pascucci, V.; Schikore, D.R. *Seed Sets and Search Structures for Optimal Isocontour Extraction*; Technical Report 99-35; Texas Institute for Computational and Applied: Austin, TX, USA, 1999.
48. Song, Y.; Liang, J.; Lu, J.; Zhao, X. An efficient instance selection algorithm for k nearest neighbor regression. *Neurocomputing* **2017**, *251*, 26–34. [[CrossRef](#)]