TOWARDS WHEAT YIELD ESTIMATION IN PLANT BREEDING FROM INHOMOGENEOUS LIDAR POINT CLOUDS USING STOCHASTIC FEATURES

T. Medic [1, *], N. Manser [2], N. Kirchgessner [3], L. Roth [3]

[1] Institute of Geodesy and Photogrammetry, ETH Zurich, Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland - tomislav.medic@geod.baug.ethz.ch

[2] Dept. of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland - nicole.manser@geo.uzh.ch

[3] Institute of Agricultural Sciences, ETH Zurich, Universitätstr. 2, 8092 Zurich, Switzerland - (norbert.kirchgessner; lukas.roth)@usys.ethz.ch

KEY WORDS: terrestrial laser scanner, TLS, agriculture, phenotyping, structural, geometry, grain yield, point cloud processing.

ABSTRACT:

The world relies heavily on wheat, corn, and rice for nutrition, with global challenges such as population growth and climate change threatening food security. To tackle this, plant breeding, supported by digital technologies, focuses on improving food quality and quantity. Currently, crop yield estimation uses indirect observations through hyperspectral data and spectral indices, such as NDVI, which suffer from low sensitivity in breeding scenarios. Terrestrial laser scanners (TLS) present an alternative, allowing observations of the quantity and morphology of wheat ears from point clouds, which are directly linked to grain yield. However, exploiting these observations under field conditions presents challenges, mainly due to reduced resolution and non-homogenous properties of point clouds. In response, we propose an approach for in-field wheat yield estimation using machine learning and stochastic features of TLS point clouds that are specifically handcrafted to be less sensitive to the abovementioned phenomena. This approach avoids the need for explicit 3D reconstruction of individual plants and plant organs. Our initial results show limited success in yield estimation when posed as a regression problem. However, when framed as a classification problem focusing on detecting top- and bottom-performing plant phenotypes, we achieved a promising accuracy of 84.4% and AUC of 0.93. While encouraging, these are only the first results under relaxed conditions and further work is needed to enhance practical applicability.

1. INTRODUCTION

Today, 50% of the calories consumed worldwide come from three main crops: wheat, corn, and rice (Awika, 2011). In the last decades, yields for wheat and rice have mainly stopped increasing, while the world population is expected to reach 9.7 billion in 2050 (United Nations, 2017). One of the ways to tackle this challenge is plant breeding, which focuses on modifying desired plant characteristics to improve the quality and quantity of food (Guo et al., 2018). Plant breeding efforts are increasingly supported by digital technologies for detecting superior plant phenotypes (phenotyping) and for estimating the main traits of interest, such as yield (Watt et al., 2020).

Currently, the estimation of crop yield by remote sensing relies on spectral data and indices such as the Normalized Difference Vegetation Index (NDVI), which indirectly infers yield from leaf properties (e.g., Ali et al., 2022). This and related approaches are well-established, however, their utility in plant breeding is limited by their sensitivity. Namely, a reported drawback of such methods is an underestimation of yields in high-yielding plants, posing the risk of overlooking high-yielding varieties, which are the focus of plant breeding (Fei et al., 2023).

There is a consensus that combining the abovementioned spectral data with plant structural parameters could alleviate this problem (Fei et al., 2023) and some research efforts demonstrated the feasibility of this hypothesis (Li et al., 2022; Sun et al., 2022). The latter studies supplemented spectral data with structural traits related to crop canopy (e.g., height, volume, and leaf area) derived from LiDAR point clouds. Their promising results were reported for phenotyping experiments where different plant varieties are treated with varying fertilization regimens, leading to substantial differences in yield among treatment groups. Hence, despite progress, the necessary sensitivity of in-field yield estimation for a normal plant breeding setting is still not achieved.

The latter studies used canopy-level 3D structural parameters, which are only indirectly related to yield. Laboratory-based experiments demonstrated that, for example, it is possible to estimate wheat grain yield with high accuracy by observing the quantity and morphological properties of wheat ears, such as length, width, and perimeter (Korohou et al., 2020). However, transferring these observations to the in-field application at a large scale carries significant challenges. One step towards this goal is estimating wheat grain yield based on wheat ear counting using high-resolution RGB images. However, this approach so far achieved only moderate success due to various obstacles, including challenging illumination conditions (Fernandez-Gallego et al. 2018). Also, such approaches fail to extract and utilize information on wheat ear morphology.

Using terrestrial laser scanners (TLS), it is possible to directly observe both wheat ear quantity and morphology (Lumme et al. 2008) without illumination-related problems. However, utilizing TLS data for in-field phenotyping poses its own challenges (Medic et al., 2023). Namely, the limited quality of the point clouds acquired in the field makes explicit 3D reconstruction of wheat ears infeasible, and inhomogeneous point cloud properties can make yield predictions more dependent on the plants' position relative to the TLS rather than on the yield itself. Hence, developing a strategy that would overcome these challenges is a prerequisite for further supporting accurate grain yield estimation using information obtained by TLSs.

* corresponding author

Figure 1. Left - Overview of the wheat variety testing field at the Eschikon Field Station of ETH Zurich; Right – Field Phenotyping Platform (FIP) carrying Faro Focus 3D S120 TLS.

To account for these challenges, we propose an approach for in-field estimation of wheat yield using stochastic features of TLS point clouds as an alternative to explicit geometrical 3D reconstruction. So far, such approaches were only used to estimate crop biomass (e.g., Xu et al., 2020) and are generally suitable only for point clouds with homogenous properties. Herein, we test the feasibility of the approaches for wheat yield estimation and present a feature engineering strategy that can reduce the impact of inhomogeneous TLS point clouds.

To demonstrate the plausibility of the approach for the *grain yield performance analysis* (our umbrella term for estimation and classification), we posed a problem both as a regression (estimation) and as a simplified classification task. Namely, the main goal of plant breeding is to detect top- and bottom-performing plant varieties to know which ones should be retained and which removed for the next breeding cycle. Hence, we demonstrate the performance of the proposed approach both on the task of explicitly estimating grain yield as well as the task of classifying top and bottom performers regarding yield.

This article is structured as follows. A short overview of the relevant state of the art, the motivation, and the goal of our study are given in this section (Section 1). Section 2 presents the implemented workflow used for grain yield regression and for classifying the top- and bottom-performing plants from TLS point clouds; Section 3 presents the experimental setup used to evaluate the implemented workflow; Section 4 presents our first results and related discussion, while the main conclusions are drawn in Section 5.

2. IMPLEMENTED WORKFLOW

The proposed workflow for TLS point cloud processing and grain yield performance analysis is established as follows. The necessary input for the algorithm are the point clouds of field plots with individual wheat varieties (see, e.g., Figure 1). For extracting the information that is related to the wheat ears' quantity and morphology, we first selected the top 40% of all points and computed per-point stochastic features (see, e.g., Weinmann et al., 2014) based on the principal component analysis (PCA) of the local point neighborhoods (3 cm radius, educated guess). Specifically, we compute:

- height (z-coordinate),
- 1st, 2nd and 3rd eigenvalue,
- number of neighbors,
- verticality,
- roughness,
- normals change rate,
- 1st order moment,
- sum of eigenvalues,

- ominivariance,
- anisotropy,
- planarity,
- linearity, and
- ratio of 1st and 2nd eigenvalue vs. sum of all eigenvalues.

As different plant organs are related to different characteristic point distributions in the point clouds (Paulus et al., 2013), we expect these stochastic features to also capture distinctive information related to wheat ears. The stochastic features extraction was implemented using the functions of CloudCompare software.

The point clouds of different field plots can have different properties depending on the relative TLS position during measurements, primarily impacting point distribution and density (see Figure 2 for the extreme case in our dataset). This effect can, in turn, bias the values of the stochastic features. Therefore, we computed normalizing features that are either less sensitive to inhomogeneous densities or can be used for quantifying the differences and accounting for them. Namely, we computed a relative point density as a ratio of point densities (number of neighbors) calculated using two neighborhood radii (3 cm and 25 cm). As additional plot-based normalizing features, we computed: per-plot point count, plot area, point number per area, mean point density, average plot height, and height difference between the top and bottom of the plot point cloud (estimated from 0.5 and 99.5 z-coordinate percentiles for robustness). Furthermore, we generated more descriptive features by computing the non-linear combinations (multiplications and divisions) of the previously computed stochastic features with the following three key features, which were present in each case: relative point density (our main normalizing feature), height, and verticality.

Such engineered features allow for the approximate isolation of points related to wheat ears, as shown in Figure 3 (top) on the example of a complex feature consisting of the following features: relative point density, verticality, height, linearity, and the sum of eigenvalues. The frequency and distribution of the feature values are related to wheat ear quantity, shape, and size. Instead of simple thresholding, segmentation, and instance counting (as exemplary presented in Figure 3, bottom), we further pursued the stochastic approach toward yield estimation and computed the discrete descriptors of the feature value distribution as the final feature engineering step.

First, we detected and removed points with extreme values using a simple 3 x MAD (median absolute deviation) threshold, i.e., we truncated eventual long tails of the feature values distribution. Then we generated histograms with 10 bins and used them to calculate the following parameters:
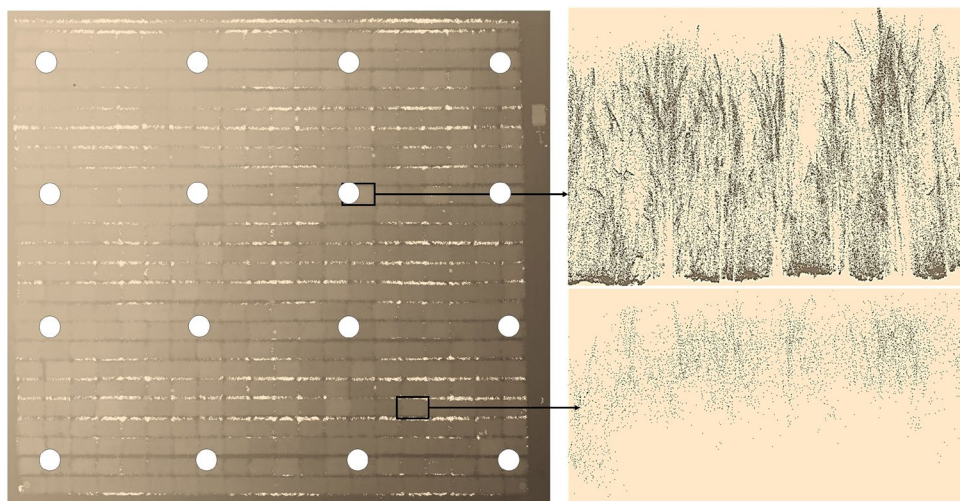
Figure 2. Left: Registered TLS point cloud of wheat variety testing field (top-down view). Right: cross-sections of individual wheat variety plots with notably different point densities - extreme cases (Up - a point cloud of a wheat plot directly below the TLS; Down – a point cloud of a wheat plot located approx. 8 m diagonally away from the TLS). Legend: ⭕ TLS station positions, terrain heights (dark – high, bright – low).

- mean feature value,
- absolute frequency of the values in the distribution tails,
- relative frequency of the values in the distribution tails,
- absolute mean values in the distribution tails,
- relative mean values in the distribution tails,
- absolute cumulative sum of the values in the distribution tails, and
- relative cumulative sum of the values in the distribution tails.

The distribution tails are defined as the following combinations of the histogram bins, where the number represents bin position from left to right: 1, 1+2, 1+2+3, 1+2+3+4, 7+8+9+10, 8+9+10, 9+10, 10. The relative frequency, mean and cumulative sum are calculated by dividing the values of the latter bin combinations by the values of the remaining bin combinations.

The described feature engineering procedure produced 2877 features in total, which were subsequently normalized using z-score normalization. As such a feature set is too large for effective training of regression and classification models using common machine learning algorithms, we added a two-step feature selection procedure. First, we computed the following custom metric for feature selection:

$$\Delta x(yield) = \frac{1}{k}\sum_{i=1}^{k} x_i^{top} - \frac{1}{k}\sum_{i=1}^{k} x_i^{bottom} \qquad (1)$$

where x denotes one of the 2877 features, subscript i denotes the field plot ($i = 1,2, ... k$), and superscripts *top* and *bottom* denote n % sub-selected plots with the highest and lowest grain yield based on the reference observations (in our study $n = 3$ %). We retained 10% of the initial features with the highest score based on Eq. 1. This selection criterion assures that the remaining features allow for separability between the top and bottom yield performers, which is the main requirement for breeding. In the second step, we used MRMR (maximum correlation, minimum redundancy) feature selection algorithm (Peng et al., 2005) to select only a set of 50 features (+ six plot-based normalizing features that were omitted from the feature selection procedure). Finally, a set of 56 features was used in the grain yield

prediction step. For the training and evaluation of both regression and classification models, we used a random 25/75 test-train split. In the process of selecting the algorithms, we examined several popular options and conducted tests on them.
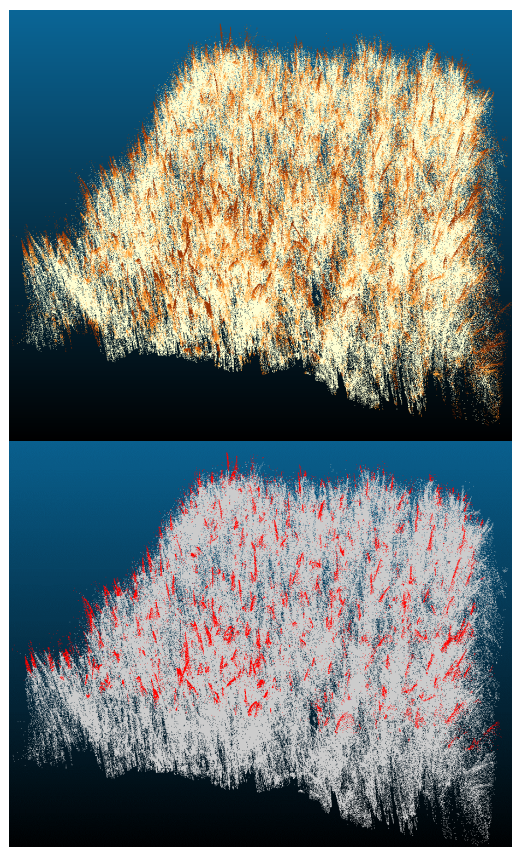

Figure 3. Point cloud of a single wheat plot colored by the values (dark - high, bright - low) of one complex handcrafted feature (combining linearity, verticality, local point density, and height) before (top) and after (bottom) simple thresholding (red – above threshold, grey – below threshold).

We tested: random forest and other bagged and boosted decision tree ensembles; support vector machines with different kernel functions (SVM), k-nearest neighbors (KNN) with different weighting functions (classification only), Gaussian Process Regression (GPR) with different kernel functions (regression only). The hyperparameters for all algorithms were tuned automatically using the Bayes optimization algorithm (Snoek et al., 2012) over 100 iterations. A 10-fold cross-validation was used to ensure a robust hyperparameter selection while preventing overfitting. The analysis was implemented in MatLab. In the results section, we present only the results of the best-performing algorithms, and the analysis is based on the standard evaluation metrics.

3. EXPERIMENT AND DATA PRE-PROCESSING

The dataset for the experimental evaluation of the workflow presented in Section 2 was collected at the Eschikon Field Station (Figure 1, left), run by the professorship of crop science at the Institute for Agricultural Science (IAS) - ETH Zurich. The point clouds were acquired using a cable-suspended field phenotyping platform (FIP) developed by the professorship (Kirchgessner et al., 2016). The platform covers an area of about 1 ha, operates from 2-5 m above the canopy, and was equipped with a Faro Focus 3D S120 TLS during the experiment (Figure 1, right).

The point clouds were acquired in the crop growing season of 2018, focusing on a winter wheat variety testing experiment consisting of 378 filed plots with dimensions of 1.7×1.4 m^2 (Figures 1 and 2, left) and hosting 360 different plant genotypes. All plots received equal treatment during the growing season, mimicking common practices in commercial wheat breeding. For more information about the agricultural practices applied within the experiment, see, e.g., Roth et al. (2020). The data was acquired with the FIP 6 times during the growing season in regular time intervals. The dataset investigated in this work was acquired on the 14th of June, shortly before the harvest.

TLS data were acquired at 16 stationary scanning positions (Figure 2, left) with a scanning resolution of 6.2 mm at a 10 m distance. The point clouds were bundled in a common (local) coordinate system using a target-based registration. Eight spherical laser-scanning targets with a radius of 10 cm were regularly distributed over the wheat variety testing field. The sphere detection and point cloud registration were done using an automatic pipeline in the manufacturer's software Faro Scene V2021.1.0.

The registered point cloud was further processed using the Open3D Python library (Zhou et al., 2018). The individual field plots were separated based on geojson files containing the 2D polygons encompassing individual plots. To avoid border effects and account for registration inaccuracy, the original plot size of 1.7×1.4 m^2 was reduced by a 25 cm buffer. This resulted in segmented plot point clouds covering an area of 1.2×0.9 m^2. The geojson files were generated using QGIS software.

Following, ground filtering was applied for each individual plot to isolate only the points related to wheat plants. This was done by computing the C2M (cloud-to-mesh) distances between the abovementioned point clouds and the TLS data of the reference epoch acquired on the 23rd of November, containing only the bare soil, i.e., the reference surface. The point cloud of the reference epoch was converted into a mesh using a Poisson

surface reconstruction algorithm implemented in Open3D – see, e.g., Becirevic et al. (2019) for more info on common practices of reference surface generation in agricultural applications. All points with the C2M distance larger than 0.1 m relative to the reference soil surface were considered as vegetation and retained for further processing.

Figure 2 (right) presents the cross-sections of two exemplary field plots with extreme differences in the point density and information content. These stark differences were the main motivations for developing the handcrafted features with a reduced sensitivity towards inhomogeneous point cloud density (Section 2). To reduce the problem complexity somewhat further, we used additional preprocessing steps to tackle the inhomogeneous point cloud density. First, we applied a spatial subsampling with a minimum point distance threshold of 2 mm, which primarily affected the wheat variety plots with extremely high point densities located directly below the TLS stations (Figure 2, right). Additionally, a subset of all 378 plots was completely removed from data processing. Namely, the plots at the low end of the point density had so poor information content that any wheat ears were barely observable (Figure 2, right). Hence, we removed 15% of the plots with the lowest point density reducing the number of plots to 320. Even after these two steps, the difference in point number between the highest and lowest density point cloud was ≈ 8.7 times (600 000 vs. 72 000 points), almost an order of magnitude.

Such pre-processed and separated point clouds of individual plots were used as the input data for grain yield performance analysis presented in Section 2. The reference measurements were acquired by manual harvesting done at the end of the growing season. The reference values are given as *grain yield* expressed in t / ha (ton per hectare).

4. RESULTS AND DISCUSSION

We initially posed the grain yield performance analysis as a regression problem, intending to establish a regression model capable of explicit grain yield estimation using the selected subset of 56 handcrafted features. The first attempt of estimating a regression model using all the 320 plot point clouds remaining after data pre-processing described in Section 3 failed. The trained models were unable to explain any portion of the variance in reference grain yield values resulting in negative or close to zero R^2 values. We presume that the following influencing factors were causing this outcome: non-negligible noise in the reference values, limited explanatory power of the handcrafted features, limited dataset size, and imbalanced dataset (small number of samples in the tails of the normally distributed reference values). We applied several strategies intending to improve the regression results (e.g., weighting the observations according to the distance from the mean value to tackle the scarcity of data at the tails of grain yield distribution), however, without notable success. Hence, these results are omitted from this work.

In the second attempt, we simplified the problem and restricted our analysis only to a subset of 320 point clouds. Namely, we separated only the top and bottom 20% of the plots with the highest and lowest grain yield, reducing the dataset to 128 point clouds. To increase the sample size for training the regression model, we implemented a data augmentation strategy. After the 25/75 test-train split, the field plots selected for training (96 plots) were segmented (cropped) from the whole point cloud (Figure 2, left) additional three times with smaller 2D polygon

sizes of 0.85×0.55 m^2 (see Section 3): once centered, once shifted 15 cm along, and once shifted 15 cm perpendicular to the sowing direction. This data augmentation strategy mimics commonly used strategies in image processing that induce image translations and cropping, and we used it to generate small but realistic variations in the extracted feature values. The training dataset, in the end, consisted of 384 point clouds.

The best regression results for the test data of this reduced subset of point clouds are presented in Figure 4. These results are attained using the SVM algorithm with the Gauss kernel, which marginally outperformed the GPR algorithm with the exponential kernel. The model exhibited some discriminative power reaching R^2 of 0.31, which is not negligible in such a challenging task as grain yield estimation in wheat breeding. The RMSE and MAE values of 2.28 and 1.80 t/ha are too high for explicit yield estimation within the required tolerance levels. However, the model was, to some degree, capable of relative ranking of the analyzed field plots considering their grain yield, which is a relevant accomplishment for plant breeding.
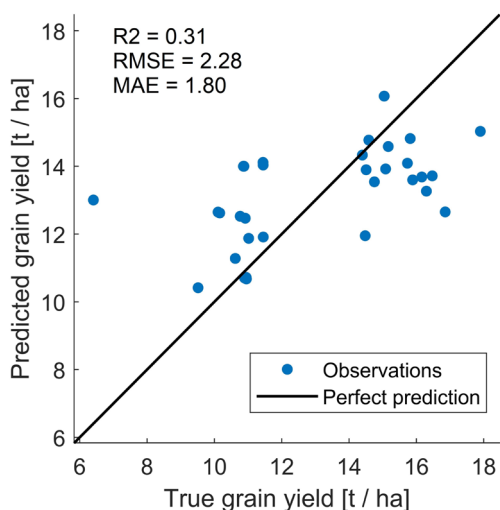

Figure 4. Regression results (test dataset): Wheat grain yield estimation based on 56 handcrafted point cloud features using a selected subset of plots present in the wheat variety testing field (top and bottom 20 % of the performers considering grain yield).

To further investigate these ranking capabilities, we transformed the grain yield estimation problem into a classification problem. Hence, we generated the ground truth class labels as: Class "Top" for the top 20 % of the plots with the highest grain yield and Class "Bottom" for the bottom 20 % of the plots. Figure 5, A) shows the classification results in the form of a confusion matrix between the top and bottom performers on a test dataset (32 plots). The results are presented for the classification with the Ensemble of Decision Trees based on Adaptive Boosting or AdaBoost (Freund and Schapire, 1995), which outperformed the other algorithms that we tested (see Section 2) on this dataset.

The implemented approach achieved a classification accuracy of 84.4 % and the AUC, the Area Under the ROC (Receiver Operating Characteristic) curve, of 0.93, which can be considered as a very good classification performance. The results achieved after repeated random test/train splits were comparable. These results are encouraging, however, they should be treated with reservation. Namely, in this simplified evaluation, we merely confirmed that it is possible to distinguish between the plots with the highest differences in the

grain yield, despite the limited spatial resolution of point clouds in the in-field phenotyping experiments and despite inhomogeneous point density. We also confirmed that the proposed handcrafted stochastic features could capture some information on the wheat ears' quantity and/or morphology.

However, in the real use case, it is necessary to solve at least a 3-class classification problem of successfully separating top, bottom, and average-performing wheat genotypes. We tested how well this extended classification task is solvable based on the proposed workflow, where we assigned three classes as follows: Class "Top" – top 20 %, Class "Middle" – middle 60 %, and Class "Bottom" – bottom 20 % performing plants considering grain yield. In this instance, the sample size was correspondingly larger (80 plots for testing and 960 plots for training after data augmentation), and the best estimation results were achieved with a different algorithm (SVM with a quadratic kernel function). The confusion matrix with the results for this classification problem is presented in Figure 5, B).
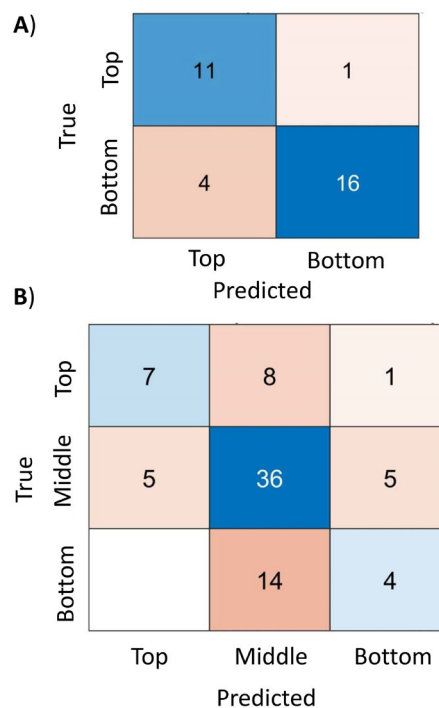

Figure 5. Classification results (test datasets): Confusion matrices for the classification task of separating wheat plots on A) top and bottom 20 % performers considering grain yield; B) top-middle-bottom performers (20-60-20 %).

The drop in the classification accuracy is significant, dropping to only 58.8 %. The AUC scores for different classes (Top: 0.75, Bottom: 0.55, Middle: 0.61) demonstrate that some discriminative power is still present, especially in the case of identifying the top-performing plants, which are of the main interest to plant breeders. Despite demonstrating some discriminative power, the presented workflow and the classification model quality need to be improved before such a procedure could be used for supporting industrial breeding procedures. Nevertheless, our analysis based on the regression and classification tasks confirmed that approaches relying on stochastic features of TLS point clouds could be used beyond biomass estimation and that they could support grain yield estimation efforts in plant breeding.

5. CONCLUSION

Within this work, we demonstrated an approach for in-field wheat yield estimation in wheat variety testing experiments using machine learning and stochastic features of TLS point clouds. The approach is designed to circumvent the need for the challenging explicit 3D reconstruction of individual plants and plant organs while still capturing information on the quantity and morphology of wheat ears. Moreover, the features were specifically handcrafted to be less sensitive to inhomogeneous TLS point cloud density, which is recognized as one of the main challenges for 3D plant phenotyping with TLS.

Our initial results, achieved with the SVM algorithm after feature and data selection, showed limited success in yield estimation: R^2 of 0.31 and MAE of 1.8 t / ha. If the problem of detecting top and bottom yield-performing wheat phenotypes is reframed as a classification problem, the results were more promising under relaxed conditions (analyzing only a subset of all field plots). We achieved an accuracy of 84.4% and AUC of 0.93 using the SVM algorithm for yield-based separation of top and bottom performing 20 % of the plots in a wheat variety testing experiment. These results were achieved despite almost an order of magnitude large differences in the point cloud densities between field plots (extremes: 600 000 vs. 72 000 points).

When the analysis was extended to all available filed plots, the results were notably poorer, indicating that the presented approach needs to be enhanced and/or integrated with other methods to achieve practical applicability. To that end, we plan to investigate the possibility of substituting a handcrafted features-based approach with end-to-end deep learning (DL), as well as combining this geometrical information extracted from TLS point clouds with the spectral information acquired with multispectral cameras. As employing DL will require solving the data scarcity problem, we will also explore a few more easily reachable options: exploring estimation methods specifically designed for predicting the tail events and exploring the information content of dynamic features derived from TLS point cloud time series.

ACKNOWLEDGMENTS

REFERENCES

Ali, A.M., Abouelghar, M., Belal, A.A., Saleh, N., Yones, M., Selim, A.I., Amin, M.E., Elwesemy, A., Kucher, D.E., Maginan, S. and Savin, I., 2022. Crop yield prediction using multi sensors remote sensing. *The Egyptian Journal of Remote Sensing and Space Science*, 25(3), pp.711-716.

Awika, J.M., 2011. Major cereal grains production and use around the world. In: *Advances in cereal science: implications to food processing and health promotion* (pp. 1-13). American Chemical Society.

Becirevic, D., Klingbeil, L., Honecker, A., Schumann, H., Rascher, U., Léon, J. and Kuhlmann, H., 2019. On the derivation of crop heights from multitemporal UAV based

imagery. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 4, pp.95-102.

Fei, S., Chen, Z., Li, L., Ma, Y. and Xiao, Y., 2023. Bayesian model averaging to improve the yield prediction in wheat breeding trials. *Agricultural and Forest Meteorology*, 328, p.109237.

Fernandez-Gallego, J.A., Kefauver, S.C., Gutiérrez, N.A., Nieto-Taladriz, M.T. and Araus, J.L., 2018. Wheat ear counting in-field conditions: high throughput and low-cost approach using RGB images. *Plant methods*, 14, pp.1-12.

Freund, Y. and Schapire, R.E., 1995. A desicion-theoretic generalization of on-line learning and an application to boosting. In: *European conference on computational learning theory* (pp. 23-37). Berlin, Heidelberg: Springer Berlin Heidelberg.

Guo, Q., Wu, F., Pang, S., Zhao, X., Chen, L., Liu, J., Xue, B., Xu, G., Li, L., Jing, H. and Chu, C., 2018. Crop 3D—a LiDAR based platform for 3D high-throughput crop phenotyping. *Science China Life Sciences*, 61, pp.328-339.

Kirchgessner, N., Liebisch, F., Yu, K., Pfeifer, J., Friedli, M., Hund, A. and Walter, A., 2016. The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system. *Functional Plant Biology*, 44(1), pp.154-168.

Korohou, T., Okinda, C., Li, H., Cao, Y., Nyalala, I., Huo, L., Potcho, M., Li, X. and Ding, Q., 2020. Wheat grain yield estimation based on image morphological properties and wheat biomass. Journal of Sensors, 2020, pp.1-11.

Kronenberg, L., Yates, S., Boer, M.P., Kirchgessner, N., Walter, A. and Hund, A., 2021. Temperature response of wheat affects final height and the timing of stem elongation under field conditions. *Journal of Experimental Botany*, 72(2), pp.700-717.

Li, Q., Jin, S., Zang, J., Wang, X., Sun, Z., Li, Z., Xu, S., Ma, Q., Su, Y., Guo, Q. and Jiang, D., 2022. Deciphering the contributions of spectral and structural data to wheat yield estimation from proximal sensing. *The Crop Journal*, 10(5), pp.1334-1345.

Lumme, J., Karjalainen, M., Kaartinen, H., Kukko, A., Hyyppä, J., Hyyppä, H., Jaakkola, A. and Kleemola, J., 2008. Terrestrial laser scanning of agricultural crops. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 37, pp.563-566.

Medic, T., Paulus, S., Bömer J., 2023. Challenges and Recommendations for 3D Plant Phenotyping in Agriculture Using Terrestrial Lasers Scanners, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. (accepted)*

Paulus, S., Dupuis, J., Mahlein, A.K. and Kuhlmann, H., 2013. Surface feature based classification of plant organs from 3D laserscanned point clouds for plant phenotyping. *BMC bioinformatics*, 14(1), pp.1-12.

Peng, H., Long, F. and Ding, C., 2005. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8), pp.1226-1238.

Roth, L., Camenzind, M., Aasen, H., Kronenberg, L., Barendregt, C., Camp, K.H., Walter, A., Kirchgessner, N. and

Hund, A., 2020. Repeated multiview imaging for estimating seedling tiller counts of wheat genotypes using drones. *Plant Phenomics*, 2020.

Snoek, J., Larochelle, H. and Adams, R.P., 2012. Practical bayesian optimization of machine learning algorithms. Advances in neural information processing systems, 25.

Sun, Z., Li, Q., Jin, S., Song, Y., Xu, S., Wang, X., Cai, J., Zhou, Q., Ge, Y., Zhang, R. and Zang, J., 2022. Simultaneous prediction of wheat yield and grain protein content using multitask deep learning from time-series proximal sensing. *Plant Phenomics*, 2022.

United Nations, and Department of Economic and Social Affairs, Population Division, 2017. World population prospects: the 2017 revision, key findings and advance tables. Working Paper No. ESA/P/WP/248 ed.

Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O. and Schurr, U., 2020. Phenotyping: new windows into the plant for breeders. *Annual review of plant biology*, 71, pp.689-712.

Weinmann, M., Jutzi, B. and Mallet, C., 2014. Semantic 3D scene interpretation: A framework combining optimal neighborhood size selection with relevant features. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*2(3), p.181.

Xu, K., Su, Y., Liu, J., Hu, T., Jin, S., Ma, Q., Zhai, Q., Wang, R., Zhang, J., Li, Y. and Liu, H., 2020. Estimation of degraded grassland aboveground biomass using machine learning methods from terrestrial laser scanning data. *Ecological Indicators*, 108, p.105747.

Zhou, Q.Y., Park, J. and Koltun, V., 2018. Open3D: A modern library for 3D data processing. *arXiv preprint* arXiv:1801.09847.