

MAPPING BARLEY LODGING WITH UAS MULTISPECTRAL IMAGERY AND MACHINE LEARNING

Odyseas Vlachopoulos ^{1,*}, Brigitte Leblon ¹, Jinfei Wang ², Ataollah Haddadi ³, Armand LaRocque ¹, Greg Patterson ³

¹ Faculty of Forestry and Environmental Management, University of New Brunswick, 2 Bailey Dr, Fredericton, NB E3B5A3 New Brunswick, Canada – (odyseas.vl, bleblon, armand.larocque)@unb.ca

² Department of Geography and Environment, University of Western Ontario, 1151 Richmond Street, ON N6A 5C2, London, Canada - jfwang@uwo.ca

³ A&L Canada Laboratories, 2136 Jetstream Rd., London, ON N5V 3P5, London, Canada - ata.haddadi@gmail.com; patterson@alcanada.com

KEY WORDS: UAS, UAV, Machine learning, Image processing, Multispectral, Precision agriculture, Random Forests, Lodging

ABSTRACT:

Unmanned Aircraft Systems (UAS) are demonstrated cost- and time-effective remote sensing platforms for precision agriculture applications and crop damage monitoring. In this study, lodging damage on barley crops has been mapped from UAS imagery that was acquired over multiple barley fields with extensive lodging damages in two aerial surveys. A Random Forests classification model was trained and tested for the discrimination of lodged barley with an overall accuracy of 99.7% on the validation dataset. The crop areas with lodging were automatically delineated by vector analysis and compared to manually delineated areas using two spatial accuracy metrics, the Area Goodness of Fit (AGoF) and the Boundary Mean Positional Error (BMPE). The average AGoF was 97.95% and the average BMPE was 0.235 m.

1. INTRODUCTION

Crop lodging is a well-known damage defined as any permanent deviation from the normal vertical position of plants and occurs mainly in barley, wheat, oat, rice, and other cereal crops. It is related to weather events such as strong wind or rain, excess fertilizers, irrigation, cultivar, soil characteristics, and disease (Pinthus, 1974; Rajkumara, 2008). Crop lodging is usually most severe at the heading and ripening growth stages. Lodging damage is responsible for significant losses in crop yield. As a result, there is a need for an accurate damage evaluation tool to determine yield losses for crop insurance paybacks. This evaluation can be done by field surveys with hand-held equipment that is assisted with Global Navigation Satellite Systems (GNSS). However, such field surveys can be laborious and time-consuming with a high probability of human errors. A suitable alternative is to use remote sensing platforms that provide high spatial resolution imagery allowing detailed crop lodging mapping. Chauhan et al. (2019) reviewed remote sensing studies on crop lodging. Space-borne platforms can be used (Chauhan et al., 2020; Shu et al., 2020), but they lack in spatial resolution and temporal flexibility due to fixed-time overpasses while being prone to weather limitations such as clouds for optical instruments. By contrast, UAS campaigns are versatile as they can be conducted at any time. They provide spatially-accurate imagery, and the data acquisition is easy and cost-efficient (Zhang and Kovacs, 2012). Chu et al. (2017) used UAS Red-Green-Blue (RGB) and Near-Infrared (NIR) imagery to map maize lodging based on plant height thresholds estimated with multivariate regression. The same method was applied on UAS RGB images for mapping barley lodging (Wilke et al., 2019). Rice lodging was mapped with UAS RGB imagery and decision tree classification (Yang et al., 2017) and with RGB and thermal infrared imagery using support vector machine (SVM) classification (Liu et al., 2018). Canola and wheat lodging was mapped with multispectral UAS imagery using SVM classification (Rajapaksa et al., 2018).

The study aims to develop a method that used multispectral MicaSense RedEdge imagery acquired from a UAS to map

lodging damages in seven barley fields during the ripening growth stage. Firstly, we compared the spectral responses of lodged and standing barley as a function of the bands. Secondly, a robust machine learning and information extraction pipeline is built, based on Random Forests and a Geographic Information Systems (GIS) vector analysis of the classification results, to automatically provide the delineated areas and boundaries of the lodging zones. The resulting areas and boundaries are then compared to manually delineated areas by using two spatial accuracy metrics, the Area Goodness of Fit (AGoF) and the Boundary Mean Positional Error (BMPE).

2. MATERIALS AND METHODS

The lodged barley study sites are in Prince Edward Island (PEI), Canada (Figure 1). Two UAS flight campaigns were conducted, Flight1 and Flight2 in late August 2018 (2018/08/20), under clear sky conditions, with a flight altitude of 120 m. Flight1 surveyed three barley fields which represent approximately 120,000m² while Flight2 surveyed four barley fields which represent approximately 170,000m². All the surveyed barley fields had extensive lodging damages (Figure 2, Figure 3).

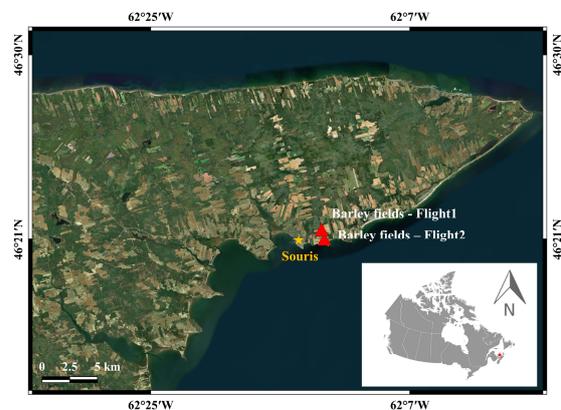


Figure 1. Location of the study sites in Prince Edward Island. ESRI Satellite (ArcGIS/World Imagery).

* Corresponding author

The UAS consisted of a DJI Matrice 100 light Unmanned Aerial Vehicle (UAV) and the MicaSense RedEdge3 multispectral camera (*MicaSense Inc., U.S.A.*) capturing data from the nadir (Table 1). The UAS MicaSense image datasets were processed using the methodology described in Figure 4. First, the images of each survey were individually inserted into a photogrammetric pipeline using the Pix4D Mapper software (*Pix4D SA, Switzerland*), performing radiometric correction, orthorectification, and stitching to produce a reflectance orthomosaic for each band.



Figure 2. Sample ground image of a barley field with lodging damage from Flight1



Figure 3. Sample ground image of a barley field with lodging damage from Flight2

Band	Blue	Green	Red	Red-Edge	NIR
Range (nm)	465–485	550–570	663–673	712–722	820–860
Bandwidth (nm)	20	20	10	10	40
Central wavelength (nm)	475	560	668	717	840

Table 1. MicaSense RedEdge band characteristics

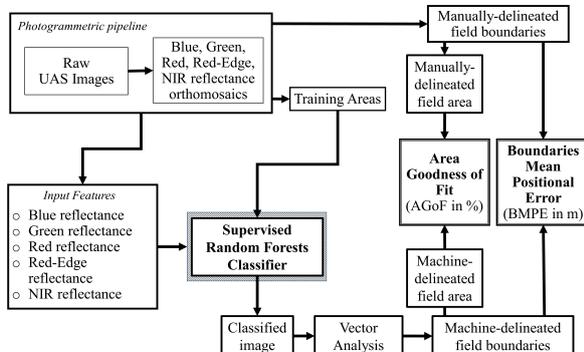


Figure 4. Flowchart of the methodology

Each orthomosaic was then inserted into a machine learning and vector analysis pipeline. The QGIS software (QGIS, 2019) was used for visualization, map generation, and other GIS and vector analysis procedures. The orthomosaics were then subjected to a supervised classification that has the three following classes: Standing Barley, Lodged Barley, and Other Vegetation. Representative training areas for each class were manually delineated and used to compute the Jeffries-Matusita (J-M) distance, which is a measure of the spectral separability of the classes. The J-M distance is the distance between random probability distributions (Richards and Jia, 2006; Wacker and Landgrebe, 1972). It has a range between 0 and 2 and it computes using the Bhattacharya (B) distance (Equations 1 and 2), which has the $[0, +\infty)$ range. For a pair of multivariate distributions, in our case a pair of classes C_1 and C_2 , the J-M distance is computed by:

$$JM_{C_1, C_2} = 2 * (1 - e^{-B_{C_1, C_2}}) \quad (1)$$

where B is the Bhattacharya distance that is computed by:

$$B_{C_1, C_2} = \frac{1}{8} * M_{C_1, C_2} + \frac{1}{2} \left\{ \log[\det(\sigma)] - \frac{\log[\det(\sigma_1)]}{2} - \frac{\log[\det(\sigma_2)]}{2} \right\} \quad (2)$$

where M is the root Mahalanobis distance (Equation 3) (Varmuza and Filzmoser, 2016) between the class means with respect to σ (Equation 4)

$$M_{C_1, C_2} = (\mu_1 - \mu_2)^t * \sigma^{-1} * (\mu_1 - \mu_2) \quad (3)$$

Where:

$$\sigma = \frac{(\sigma_1 + \sigma_2)}{2} \quad (4)$$

With:

μ_1, μ_2 = mean of classes C_1 and C_2 , respectively

σ_1, σ_2 = covariance matrices for classes C_1 and C_2 , respectively

The J-M lowest value of 0 indicates that the distributions are highly correlated, thus the classes are spectrally not separable, whereas the upper asymptotic limit of 2 indicates excellent class spectral separability. Since the J-M distance assesses the separability of the training datasets before classification, high J-M distances indicate that photointerpretation and in-situ human perception of lodging damage are good enough for the machine learning pipeline.

The training areas were then used in a supervised classifier, Random Forests (RF). RF was introduced by Breiman (Breiman, 2001, 2003) and implemented in R (Liaw and Wiener, 2002). RF has been widely used in remote sensing (Belgiu and Drăguț, 2016) due to its demonstrated statistical robustness and deployment advantages in parametrization and computational needs. RF is a non-parametric decision tree ensemble supervised classification model that does not assume a normal distribution of the data. In our processing, we set the number of trees grown (*ntree*) at 500 and we kept the number of random features selected for the growth of each tree from the feature space (*mtry*) at its default value, which is the square root of the size of the feature space used, rounded down. The classification accuracy is determined through the RF Out-Of-Bag (OOB) error rate, which is the aggregation of individual OOB error rates from each individual tree grown, when the data omitted (~37%) from the training of the decision trees due to the random bootstrapping with replacement strategy in the data

sampling are parsed and classified by the trees that were not trained with them. The OOB error rate is the complementary percentage of the overall classification accuracy. The RF OOB confusion matrix allows the computation of the class user's and producer's accuracies (Congalton, 1991). The RF implementation in R generates two feature importance matrices: 1) the *MeanDecreaseGini*, which represents the mean decrease in the Gini index of node impurity when a feature is split at the creation of a node and 2) the *MeanDecreaseAccuracy*, which represents the mean decrease of prediction accuracy when a feature is permuted. For the graphical representations and plots in this study, the *ggplot2* (Wickham, 2016) library was used.

The resulting classified images were afterward cleaned from some salt and pepper noise using a 1-pixel radius Majority filter of the System for Automated Geoscientific Analyses (SAGA) (Conrad et al., 2015). The resulting classified images were then inserted into the vector analysis pipeline. The *polygonize* function of the GDAL library (GDAL/OGR, 2019) was employed to vectorize the classified fields, and the results were refined by disposing minor patches of mixed lodged and standing barley within the damaged areas and finally smoothing the borders with equal vector buffering and debuffering. The resulting lodged areas were compared to lodged areas that were manually delineated over the UAS orthomosaics through photo-interpretation. Such delineation was made easier thanks to the very high spatial resolution of the data (~8cm) and the geolocation accuracy of the UAS data and the photogrammetric procedure in the making of the final orthomosaics. Following Vlachopoulos et al. (2020a, 2020b), the accuracy of the methodology was estimated by two spatial accuracy metrics: the Area Goodness of Fit (AGoF) and the Boundary Mean Positional Error (BMPE). AGoF computes the percentage of area overlap between the manual and the machine learning derived polygon areas as follows (Equation 5).

$$AGoF = \left(\frac{C}{AC + C} \right) * \left(\frac{C}{BC + C} \right) \quad (5)$$

Where:

A = manual delineated polygon area (m²)

B = machine delineated polygon area (m²)

C = manual and machine delineated polygons intersection area (m²)

AC = | A - C |

BC = | B - C |

BMPE calculates the aggregated minimum distances between the manually and the geographical points sampled at 0.5m on the machine delineated polygons (Equation 6). The smaller the BMPE output, the higher the overall positioning accuracy.

$$BMPE = \frac{1}{N} \sum_{i=1}^N MinDist_i \quad (6)$$

Where:

N = number of point samples

MinDist_i = minimum distance between the ith point sample and the manually delineated polygons

3. RESULTS AND DISCUSSION

Table 2 presents the J-M distances between each pair of classes computed with all the MicaSense bands and the training data related to the three classes (Other Vegetation, Standing Barley, and Lodged Barley). The resulting distances indicate excellent class spectral separabilities for every pair of classes, given that all the J-M distances are greater than 1.9.

Class pair	J-M distance
Other vegetation - Standing Barley	1.966885
Other vegetation - Lodged Barley	1.995820
Standing Barley - Lodged Barley	1.915879

Table 2. Jeffries-Matusita distance for every pair of classes computed with all the MicaSense bands

We trained and tested an RF classification model over approximately 50,000 randomly chosen pixels for each class using spatially dispersed training sites from Flight1 fields. This method achieves a spatially robust and minimally biased training set for the machine learning algorithm. The RF OOB error rate was 0.65% and the overall accuracy was 99.35%. The detailed confusion matrix with the User's Accuracy (UA), the Error of Commission (EC), the Producer's Accuracy (PA), and the Error of Omission (EO) is shown in Table 3.

Class	Other Vegetation	Standing Barley	Lodged Barley	UA (%)	EC (%)
Other Vegetation	49987	24	0	99.95	0.05
Standing Barley	30	49646	358	99.22	0.78
Lodged Barley	0	567	49458	98.87	1.13
PA (%)	99.94	98.82	99.28		
EO (%)	0.06	1.18	0.72		

Table 3. Random Forests OOB confusion matrix with User's and Producer's Accuracies and Errors of Omission and Commission. The bold figures indicate correctly classified pixels for each class.

The RF model was finally assessed for its accuracy and transferability with a validation dataset, spatially independent from the training and testing dataset, using only data from Flight2 fields. Approximately 20,000 randomly chosen pixels from spatially dispersed training sites for each class were used for validation, with an overall accuracy of 99.7%. The related confusion matrix with UA, EC, PA and EO are shown in Table 4.

Class	Other Vegetation	Standing Barley	Lodged Barley	UA (%)	EC (%)
Other Vegetation	18571	1	0	99.99	0.01
Standing Barley	3	19850	37	99.8	0.2
Lodged Barley	1	157	19973	99.22	0.78
PA (%)	99.98	99.21	99.82		
EO (%)	0.02	0.79	0.18		

Table 4. Random Forests confusion matrix for the validation dataset with User's and Producer's class Accuracies and Errors of Omission and Commission. The bold figures indicate correctly classified pixels for each class.

Our overall validation accuracy of 99.7% is higher than the one of Yang et al. (2017), who used a decision tree classification on UAV RGB imagery to map lodged rice crops (96.17%), and Rajapaksa et al. (2018), who applied the SVM classification to multispectral images to map canola (96%) and wheat (92.6%) crops.

Figure 5 shows the *MeanDecreaseAccuracy* and *MeanDecreaseGini* feature importance plots for the classification produced by RF. For both metrics, the Red reflectance is the dominant feature. The red reflectance is also the one that allows the highest graphical discrimination between the three classes as shown by the boxplots of the reflectance values computed with the training pixels of each class and each band (Figure 6).

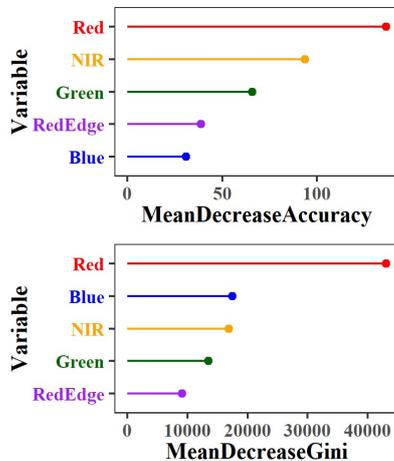


Figure 5. Random Forest *MeanDecreaseAccuracy* and *MeanDecreaseGini* feature importance metrics

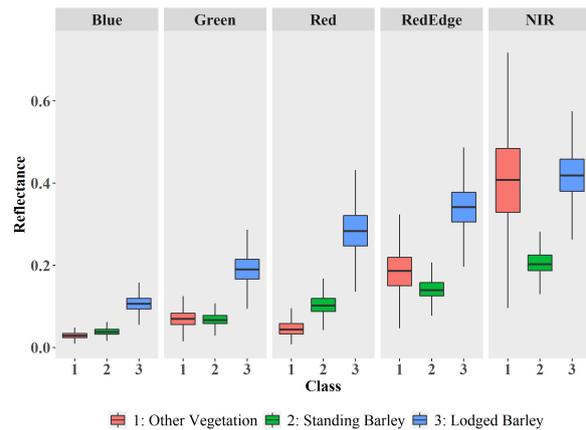


Figure 6. Boxplots of the mean reflectance values of each class extracted from the training areas

Figure 7 presents the spectral signatures of the three classes for the five multispectral bands. The plot shows a quite clear discriminating signature of lodged barley and the rest of the classes for all the bands except the NIR. Both Figure 6 and Figure 7 suggest that the Lodged Barley class can be easily discriminated using the Micasense bands as feature space. All reflectances show on average very higher values for lodged barley compared to the other vegetation and the standing healthy barley. Compared to the other vegetation which is green, the lodged barley which is yellow shows very high reflectance in the red due to the loss of chlorophyll. With respect to the healthy standing barley, the reflectance difference is probably due to a high amount of shadow in the healthy standing barley, which decreases the overall reflectance level.

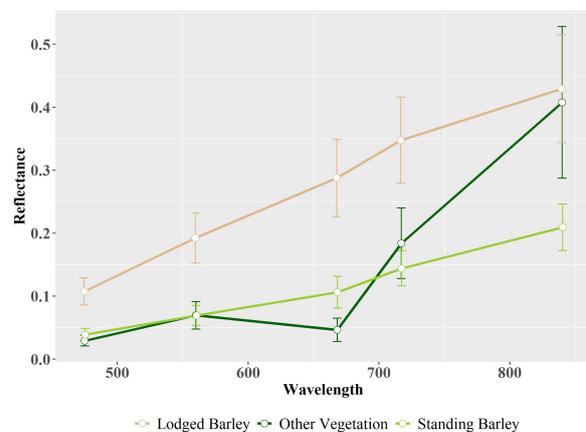


Figure 7. Spectral signatures of the three classes

The classifier was trained and validated upon the fields from the first flight while the second flight barley fields were exclusively used to validate the generated RF classification. Both flight fields were used to assess the spatial accuracy of lodging damage assessment after the vectorization procedure with the AGoF and BMPE metrics. The RGB true color composite and the resulting classified image for the Flight1 fields are shown in Figure 8 and for Flight2 fields in Figure 9. The vector analysis for Flight1 resulted in an AGoF of 96.8% and a BMPE of 0.17m. For Flight2, the AGoF was 99.1% and the BMPE was 0.3m.

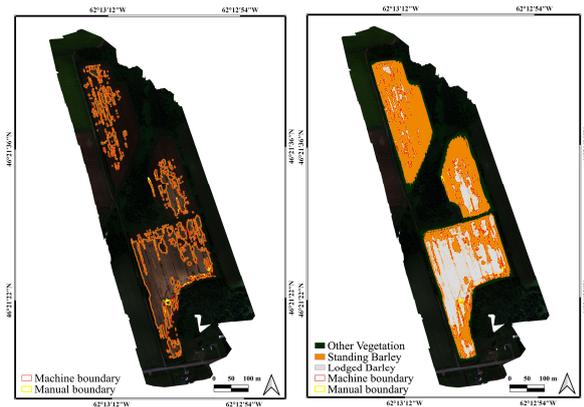


Figure 8. RGB composite and classification image of Flight1 barley fields with the lodging delineation results

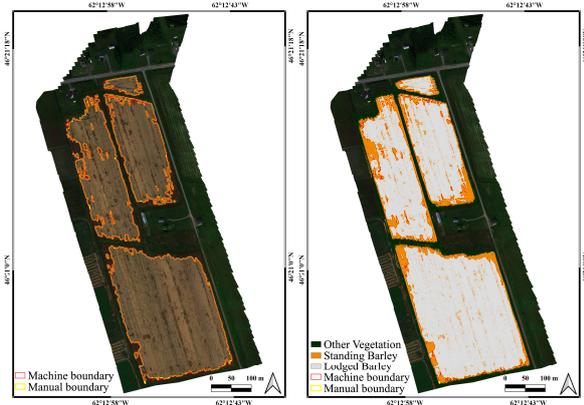


Figure 9. RGB composite and classification image of Flight2 barley fields with the lodging delineation results

4. CONCLUSIONS

Two flight campaigns were conducted for surveying seven barley fields with lodging damage during the ripening growth stage in late August 2018 in PEI, Canada, using a UAS equipped with a multispectral MicaSense RedEdge camera. Firstly, a photogrammetric pipeline was utilized to produce the orthomosaics from each flight. The orthomosaics were inserted into a machine learning pipeline that used the RF classification model followed by a vector analysis for the delineation of the damaged barley areas from the classified images of each flight. The overall accuracy of the RF classification was greater than 99%, showing that the RF classifier is a robust and consistent machine learning tool for the classification of UAS imagery in three classes (standing barley, lodged barley, other vegetation).

The vector analysis produced automatically delineated areas and borders were compared to a photo-interpreted delineation of the damaged areas. The comparison produced excellent results for both flights and all barley fields, with an average AGoF of 98% and an average BMPE of 0.23m. Such results can enable the timely and accurate estimation of the damages on barley crops for yield estimation and insurance purposes. The proposed methodology was limited to the classification of ripening lodged barley and similar studies should be undertaken to test the method over lodging damages in other growth stages for barley and on other crops. Additional information derived from spectral data such as textural and vegetation indices should be tested in future work, especially when surveying lodging

damage on different crops and growth stages. Such applications will allow early detection of crop lodging and strengthen agricultural decision support systems for possibly preventing more extensive damages.

ACKNOWLEDGMENTS

This research was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC Canada) grant number CRDPJ507141-16 awarded to Prof. Dr. Brigitte Leblon (University of New Brunswick) and Prof. Dr. Jinfei Wang (University of Western Ontario). The images were acquired under the NRC-IRAP Project number 927849 awarded to A&L Canada Laboratories.

REFERENCES

- Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 24–31.
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., 2003. Manual for setting up, using, and understanding Random Forest V4.0.
- Chauhan, S., Darvishzadeh, R., Boschetti, M., Pepe, M., Nelson, A., 2019. Remote sensing-based crop lodging assessment: Current status and perspectives. *ISPRS Journal of Photogrammetry and Remote Sensing* 151, 124–140. <https://doi.org/10.1016/j.isprsjprs.2019.03.005>
- Chauhan, S., Darvishzadeh, R., Lu, Y., Boschetti, M., Nelson, A., 2020. Understanding wheat lodging using multi-temporal Sentinel-1 and Sentinel-2 data. *Remote Sensing of Environment* 243, 111804. <https://doi.org/10.1016/j.rse.2020.111804>
- Chu, T., Starek, M.J., Brewer, M.J., Murray, S.C., Pruter, L.S., 2017. Assessing lodging severity over an experimental maize (*Zea mays* L.) field using UAS images. *Remote Sensing* 9, 923. <https://doi.org/10.3390/rs9090923>
- Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37, 35–46. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B)
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1. 4. *Geoscientific Model Development* 8, 1991–2007.
- GDAL/OGR contributors (2019). GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation. <https://gdal.org>
- Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R news* 2, 18–22.
- Liu, T., Li, R., Zhong, X., Jiang, M., Jin, X., Zhou, P., Liu, S., Sun, C., Guo, W., 2018. Estimates of rice lodging using indices derived from UAV visible and thermal infrared images. *Agricultural and Forest Meteorology* 252, 144–154. <https://doi.org/10.1016/j.agrformet.2018.01.021>

- Pinthus, M.J., 1974. Lodging in wheat, barley, and oats: the phenomenon, its causes, and preventive measures, in *Advances in Agronomy*. Elsevier, 209–263. *Precision Agriculture* 13, 693–712. <https://doi.org/10.1007/s11119-012-9274-5>
- QGIS Development Team (2019). QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>
- Rajapaksa, S., Eramian, M., Duddu, H., Wang, M., Shirliffe, S., Ryu, S., Josuttis, A., Zhang, T., Vail, S., Pozniak, C., Parkin, I., 2018. Classification of crop lodging with gray level co-occurrence matrix. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 251–258. <https://doi.org/10.1109/WACV.2018.00034>
- Rajkumara, S., 2008. Lodging in cereals—A review. *Agricultural reviews* 29, 55–60.
- Richards, J.A., Jia, X., 2006: *Remote Sensing Digital Image Analysis: An Introduction, 4th ed.* Springer-Verlag, Berlin Heidelberg.
- Shu, M., Zhou, L., Gu, X., Ma, Y., Sun, Q., Yang, G., Zhou, C., 2020. Monitoring of maize lodging using multi-temporal Sentinel-1 SAR data. *Advances in Space Research* 65, 470–480. <https://doi.org/10.1016/j.asr.2019.09.034>
- Varmuza, K., Filzmoser, P., 2016: *Introduction to Multivariate Statistical Analysis in Chemometrics*. CRC Press. <https://doi.org/10.1201/9781420059496>
- Vlachopoulos, O., Leblon, B., Wang, J., Haddadi, A., LaRocque, A., Patterson, G., 2020a. Delineation of crop field areas and boundaries from UAS imagery using PBIA and GEOBIA with Random Forest classification. *Remote Sensing* 12, 2640. <https://doi.org/10.3390/rs12162640>
- Vlachopoulos, O., Leblon, B., Wang, J., Haddadi, A., LaRocque, A., Patterson, G., 2020. Delineation of Bare Soil Field areas from unmanned aircraft system imagery with the mean shift unsupervised clustering and the Random Forest supervised classification. *Canadian Journal of Remote Sensing* 46, 489–500. <https://doi.org/10.1080/07038992.2020.1763789>
- Wacker, A.G., Landgrebe, D.A., 1972. Minimum distance classification in remote sensing. *LARS Technical Reports* 25.
- Wickham, H., 2016. *ggplot2: elegant graphics for data analysis*. Springer.
- Wilke, N., Siegmann, B., Klingbeil, L., Burkart, A., Kraska, T., Muller, O., van Doorn, A., Heinemann, S., Rascher, U., 2019. Quantifying Lodging Percentage and Lodging Severity Using a UAV-Based canopy height model combined with an objective threshold approach. *Remote Sensing* 11, 515. <https://doi.org/10.3390/rs11050515>
- Yang, M.-D., Huang, K.-S., Kuo, Y.-H., Tsai, H.P., Lin, L.-M., 2017. Spatial and spectral hybrid image classification for rice lodging assessment through UAV imagery. *Remote Sensing* 9, 583. <https://doi.org/10.3390/rs9060583>
- Zhang, C., Kovacs, J.M., 2012. The application of small unmanned aerial systems for precision agriculture: a review.