agronomy

MDPI

Article

Prediction of Turfgrass Quality Using Multispectral UAV Imagery and Ordinal Forests: Validation Using a Fuzzy Approach

Alexander Hernandez [1,*], Shaun Bushman [1], Paul Johnson [2], Matthew D. Robbins [1] and Kaden Patten [1]

[1] Forage and Range Research Laboratory, Agricultural Research Service, United States Department of Agriculture, Logan, UT 84322, USA; shaun.bushman@usda.gov (S.B.); matthew.robbins@usda.gov (M.D.R.); kaden.patten@usda.gov (K.P.)
[2] Plants, Soils and Climate Department, Utah State University, Logan, UT 84322, USA; paul.johnson@usu.edu
* Correspondence: alexander.hernandez@usda.gov; Tel.: +1-4355355972

Abstract: Protocols to evaluate turfgrass quality rely on visual ratings that, depending on the rater's expertise, can be subjective and susceptible to positive and negative drifts. We developed seasonal (spring, summer and fall) as well as inter-seasonal machine learning predictive models of turfgrass quality using multispectral and thermal imagery collected using unmanned aerial vehicles for two years as a proof-of-concept. We chose ordinal regression to develop the models instead of conventional classification to account for the ranked nature of the turfgrass quality assessments. We implemented a fuzzy correction of the resulting confusion matrices to ameliorate the probable drift of the field-based visual ratings. The best seasonal predictions were rendered by the fall (multi-class AUC: 0.774, original kappa 0.139, corrected kappa: 0.707) model. However, the best overall predictions were obtained when observation across seasons and years were used for model fitting (multi-class AUC: 0.872, original kappa 0.365, corrected kappa: 0.872), clearly highlighting the need to integrate inter-seasonal variability to enhance models' accuracies. Vegetation indices such as the NDVI, GNDVI, RVI, CGI and the thermal band can render as much information as a full array of predictors. Our protocol for modeling turfgrass quality can be followed to develop a library of predictive models that can be used in different settings where turfgrass quality ratings are needed.

Keywords: turfgrass quality; multispectral imagery; unmanned aerial vehicles; ordinal forests; fuzzy corrections; remote sensing; predictive modeling

Citation: Hernandez, A.; Bushman, S.; Johnson, P.; Robbins, M.D.; Patten, K. Prediction of Turfgrass Quality Using Multispectral UAV Imagery and Ordinal Forests: Validation Using a Fuzzy Approach. *Agronomy* **2024**, *14*, 2575. https://doi.org/10.3390/agronomy14112575

Academic Editor: Baohua Zhang

Received: 25 September 2024
Revised: 26 October 2024
Accepted: 29 October 2024
Published: 1 November 2024

1. Introduction

The development of turfgrasses that have low water requirements is important as they play significant roles in urban landscaping, recreational areas, erosion control, and other applications with increasing demands for reductions in water use [1,2]. Oftentimes, turf grasses are exposed to environmental stresses, particularly drought [3]. Drought stress can have severe impacts on turf quality, causing wilting, browning, and even death of the grass, affecting both aesthetics and functionality. Research has demonstrated that selecting drought-resistant turfgrass cultivars can effectively reduce water use while maintaining turf quality [4]. Enhancing the drought resistance of turf grasses is essential to ensure their sustainability and functionality under water-limited conditions. Research has underscored the importance of improving drought tolerance in cool-season turfgrasses through targeted breeding efforts and precise drought screening protocols [5]. Identifying and screening grasses with excellent drought tolerance is a desirable strategy in breeding drought-tolerant turf cultivars [6]. This highlights the necessity for ongoing research and development to boost the resilience of turf grasses to water scarcity.

Traditionally, turfgrass quality assessment has relied on visual methods, such as the National Turfgrass Evaluation Program's (NTEP) 1 to 9 scale [7], where 1 indicates low-quality turf, 9 indicates high-quality turf, and 6 is considered the minimum acceptable level

of turfgrass quality [8]. This visual rating system has been widely adopted by researchers and turfgrass managers globally for evaluating turfgrasses [9]. Assessing turf quality through visual ratings offers both advantages and disadvantages. Visual ratings are a common and cost-effective method used in the turfgrass industry to evaluate the overall quality of turf based on parameters such as color, density, and uniformity [10]. This method provides a quick and straightforward way to assess the aesthetic appeal and health of turfgrass [11]. Visual ratings are subjective but can still offer valuable insights into the condition of the turf, making them accessible to a wide range of users [12]. One of the key advantages of using visual ratings is the simplicity and ease of implementation. It does not require specialized equipment, making it a practical choice for routine assessments of turf quality. Additionally, visual ratings can provide immediate feedback on the condition of the turf, allowing for timely interventions to maintain or improve turf quality [13,14]. This method is also versatile, as it can be applied at different frequencies, such as weekly, bi-weekly, or monthly, depending on the study or management needs [15]. However, there are also limitations to assessing turf quality through visual ratings. One of the main drawbacks is the subjectivity involved in the assessment process. Different individuals may perceive turf quality differently, leading to potential inconsistencies in ratings. This subjectivity can be influenced by factors such as lighting conditions, observer experience, and personal bias, which may affect the reliability of the results. Moreover, visual ratings may not capture subtle changes in turf quality that could be detected through more objective methods like digital image analysis [9].

In recent years, the use of unmanned aerial vehicles (UAVs) and digital photogrammetry has revolutionized the monitoring and assessment of various environmental features, including turfgrass quality. The high-resolution outputs generated through drone imagery provide an unprecedented opportunity to move beyond traditional classification systems and work with spatially explicit continuums of data [16]. Moreover, the integration of digital image analysis and spectral reflectance techniques has been instrumental in determining turfgrass quality. In a study in Kansas, USA, canopy spectral reflectance, which can be measured using normalized difference vegetation index (NDVI), proved to offer an objective means to evaluate the visual quality of turfgrass [17]. The results of this study indicated, however, that NDVI can provide precise predictions in plots maintained at the same mowing height and in experiments with the same species. In another study conducted in Las Cruces, New Mexico, USA, NDVI was found to be strongly associated with turfgrass quality, outperforming other digital image analysis outputs in assessing parameters like a dark green color index or percent green turfgrass coverage [9]. This study emphasized that replacing visual assessments with digital image analysis may be questionable because visual ratings can detect aesthetic appeal differences between varieties much more accurately. Research conducted in northern Italy showed how multispectral radiometry and vegetation indices like NDVI can provide quantitative and objective evaluations of turfgrass quality responses to various stresses [18], further strengthening the idea that NDVI may be used to reduce the subjective nature of visual quality assessments.

The aforementioned techniques enable a more comprehensive and accurate assessment of turfgrass quality compared to traditional visual methods such as the NTEP. Furthermore, the use of UAVs in conjunction with thermal and multispectral imagery has shown promise in estimating actual evapotranspiration over irrigated turfgrass surfaces. Studies have highlighted the efficacy of models like the Two-Source Energy Balance (TSEB) model in utilizing drone imagery to estimate spatial variations of daily actual evapotranspiration. The TSEB is valuable for landscape irrigation management, especially under drought conditions [19]. The combination of drone technology and advanced modeling techniques enhances the precision and efficiency of assessing turfgrass quality parameters related to water management and environmental conditions.

Here we report on the development of a novel method to predict turf grass quality using imagery collected by multispectral sensors onboard UAVs over the span of two years (2022 and 2023) over 300 plots containing multiple varieties of a widely used turf grass:

Kentucky Bluegrass (KBG; *Poa pratensis*). Our objectives were twofold: (a) evaluate the feasibility of fitting seasonal and global models to predict turfgrass quality in semiarid landscapes, and (b) quantify the impact of implementing fuzzy accuracy assessment approaches on the overall model fit. We expect that the development of transparent and repeatable models such as the ones presented in this paper will gradually contribute to obtaining more objective predictions of turf quality across different grass species, ecological conditions, and other physical modifiers across the country.

2. Materials and Methods

2.1. Study Area

This study was conducted at the Utah State University (USU) Greenville Research Farm (41°45′56.41″ N, 111°48′37.42″ W) in Logan, UT, USA. The site is at 1409 m above sea level with a mean daily temperature of 9.2 °C and an average annual precipitation of 479 mm [20]. The experiment included 100 different KGB varieties, replicated three times in a 1 m × 1 m plot, totaling 300 units. This experiment uniformly irrigated weekly throughout the season. The plots' size and spatial arrangement, as well as the distances between repetitions of this field experiment, followed the NTEP guidelines. Figure 1 shows the location of the experiment as well as the distribution of the KBG varieties.

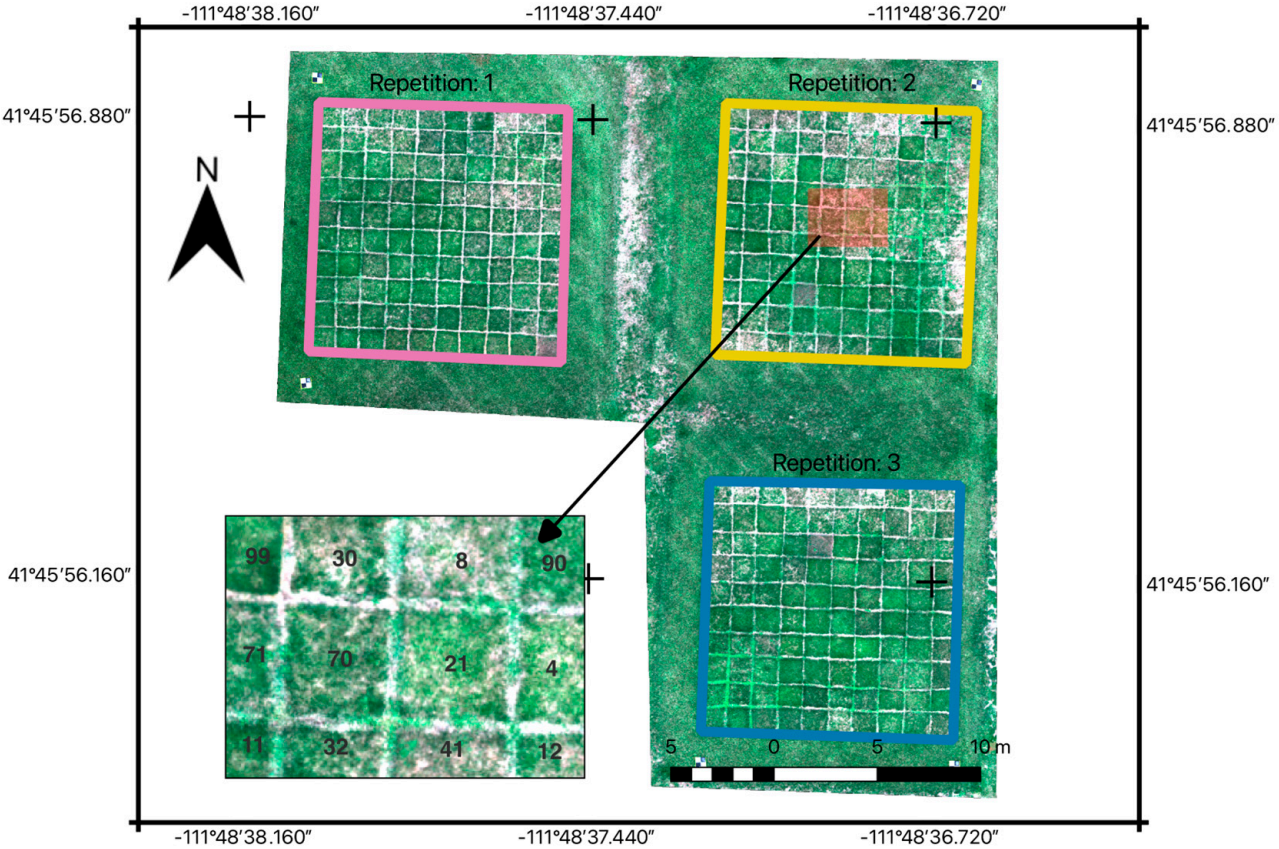

Figure 1. Layout of the repetitions and experimental units used in this study, overlaid on natural color UAV images collected on 16 August 2022. Inset map provides a close-up representation of the KBG plots and their randomization within each repetition.

2.2. Collection of Turf Grass Quality Ratings

Each one of the 300 plots was visually assessed and given a quality rating (from 1 to 9) using the standardized NTEP norms [7] on the same dates as UAV imagery acquisition (Table 1). Since quality ratings and UAV imagery needed to be acquired on the same day for predictive modeling, the selection of the dates was dictated by weather conditions optimal for UAV flights outlined below.

Table 1. List of dates and associated season when the 300 plots were rated using the NTEP methodology.

Year	Date	Season
2022	28 July	Summer
2022	16 August	Summer
2022	20 September	Fall
2023	28 April	Spring
2023	16 May	Spring
2023	30 June	Summer
2023	31 August	Summer
2023	25 September	Fall
2023	23 October	Fall

The ratings were conducted on clear days and were performed by three turfgrass quality raters from both USU and the US. Department of Agriculture, Agricultural Research Service (USDA/ARS) Forage and Range Research Lab (FRRL) personnel with proper training on the NTEP norms to assign turf grass quality. To collect the best quality imagery, the UAV must be flown within one hour of solar noon, under clear and sunny days, and when winds do not exceed the aircraft's maximum tolerances. The ratings were conducted during the dates shown above to track changes between seasons and to explore if seasonal models would perform differently. The selection of the specific dates was arbitrary and depended on a combination of clear weather conditions, visual raters' availability, and resource logistics. The start of the project was contemplated to be in the spring of 2022. However, numerous logistic difficulties (i.e., inability to fully integrate UAV and multispectral sensors and insufficient UAV pilot training) prevented us from collecting data during that season. Once these limitations were overcome, data collection formally started in the summer of 2022 until the fall of 2023.

2.3. UAV Imagery Acquisition

A Matrice 600 Pro (Shenzhen DJI Sciences and Technologies Ltd., Shenzhen, Guangdong, China) hexacopter UAV was used, with a multispectral Micasense Altum (AgEagle Aerial Systems Inc., Wichita, KS, USA) sensor. The Altum collected imagery in six bands of the electromagnetic spectrum. A sample schematic of the research plots and the UAV utilized during this research is in Figure 2.

Figure 2. The KBG (1 m × 1 m) plots at the Greenville Farm and the UAV carrying the multispectral sensor during one typical clear day of data collection.

The band name designation, along with the center wavelength for this sensor, is provided in Table 2.

Table 2. Band names, center wavelength and bandwidth of the Altum sensor used in the study.

Band Name and File Order [1]	Center Wavelength (nm [2])	Band Amplitude (nm)
Blue **B**—1	475	32
Green **G**—2	560	27
Red **R**—3	668	14
Red Edge **RE**—4	717	12
Near Infrared **NIR**—5	842	57
Longwave Infrared **LWIR**—6	11,000 (11.0 μ)	6000 (6.0 μ)

[1] Letters in bold are abbreviations to be used throughout the text. Numbers indicate the sequence in which band files are stored. [2] nm—nanometers~$^{1}/_{1000}$ microns μ.

Flight missions for each date (Table 1) were prepared using the UgCS software (https://www.sphengineering.com/flight-planning/ugcs, accessed on 10 May 2022) version 5.3.0. Missions were flown at an altitude of 34.8 m, which rendered a pixel resolution of 1.5 cm for the multispectral bands (B, G, R, RE, and NIR) and 23.2 cm for the LWIR. The missions were planned so that the flight lines were flown from east to west and west to east, and the speed of the UAV was kept consistent to ensure that image frames had a minimum overlap (front and side) of 75%. Immediately before the launch and at landing, image frames of the calibrated reflectance panels were collected. This was done to create reflectance-compensated image outputs during post-processing. All the flights were conducted during solar noon, under clear and sunny conditions with low wind speeds. The sensor always pointing at the nadir (i.e., perpendicular to the ground). Ground control points (GCPs) were placed strategically throughout the flight mission area for georectification of orthomosaics during image processing (see below). The GCPs' real-world coordinates (i.e., Latitude, Longitude and elevation) were collected using Emlid RS2+ (https://emlid.com/reachrs2plus/, accessed on 3 March 2022) receivers.

2.4. UAV Imagery Processing

2.4.1. Conversion to Reflectance

We used a modified version of the Micasense image processing scripts (https://github.com/micasense/imageprocessing, accessed on 10 October 2023) that utilized the calibrated reflectance panels with accompanying imagery to control the radiometric workflow that converts raw digital numbers into reflectance information. In addition, the different bands or files (Table 2) are co-registered such that features in the imagery are geometrically aligned from band to band. The conversion to reflectance was carried out for the B, G, R, RE, and NIR bands. Since the information collected on the LWIR band is of emitted (not reflected) radiation, there was no radiometric conversion, and the algorithm ensured that the LWIR bands were co-aligned with other reflective bands. The Micasense scripts were modified to (1) utilize the ORB [21] algorithm instead of the SIFT [22] algorithm for feature detection, decreasing processing time per capture by a factor of 10; (2) implement parallel processing to reduce processing time; (3) streamline the workflow into a single wrapper script executed by command as opposed to Jupyter notebooks; and (4) output individual band layers, the required downstream input, instead of stacked images.

2.4.2. Stitching the Images into Georectified Orthophotomaps

Individual reflectance image frames were stitched together into orthophotos using the Open Drone Map (ODM) software version 3.1.7 [23]. The GCPs collected at the field were used as reference points within the ODM workflow to generate orthomosaics that were georectified or geometrically corrected to a spatial grid—Universal Transverse Mercator UTM Zone 12 North. Orthomosaics that have been georectified using the GCPs allow for a full correspondence or digital alignment of an aerial image with other images taken over the same area or with other maps that are in the same spatial grid.

2.4.3. Extraction of Vegetation Indices (VIs)

We calculated several different VIs (Supplementary Material—Table S1) using custom Python scripts from the light reflectance intensities detected by the multispectral bands of the Altum sensor (Table 2). The single-band TIFF files were read in as arrays using the *Rasterio* version 1.3.9 [24] package, and the calculations were done using the *GeoWombat* version 2.1.0 [25] package. The *GeoWombat* package allows on-the-fly geographic transformations, automatic data alignment, and flexible data writing over parallel tasks. Using the *GeoWombat* package significantly improved the calculation speeds of the VIs.

2.5. Preparation of the Modeling Matrix

2.5.1. Zonal Statistics

To simplify our analysis, we computed zonal statistics, which are representative values (i.e., mean, median, mode, etc.) for all the pixels that are completely contained within a plot's polygon. With a ground sample distance of 1.5 cm and an average plot size of 1 m^2, approximately 4445 pixels per plot were captured. We focused on median values and extracted zonal statistics for all the remote sensing bands (B, G, R, RE, NIR, and LWIR) and all the subsequently derived vegetation indices (i.e., NDVI, RVI, NDRE). For this process, we used the Python package *exactextract* version 0.2.0 [26], which accounts for partial coverage of pixels. This package is rapid, reduces computational time, and provides more accurate estimates.

2.5.2. Modeling Matrix

We prepared a matrix for modeling purposes where the rows represent the individual observations (300 KGB plots × 9 visual ratings (Table 1) for each plot for a total of 2700 rows or observations), and the columns can be described as (a) the response variable that was to be predicted: visual ratings with values ranging from 1–9 as previously (Section 2.1) described, and (b) the median value (i.e., zonal statistics) for all the bands and vegetation indices that were derived.

2.6. Modeling

2.6.1. Selection of the Modeling Algorithm

The response variable values (1–9 turf quality ratings) follow an incremental order such that plots with a visual rating of 1 are of lesser quality than the plots with a visual rating of 2, plots with a visual rating of 2 are of lesser quality than the plots with a visual rating of 3, and so forth. In this context, we conducted the modeling of the response variable using ordinal regression (OR) [27]. OR can manage multiple ordered categories in the response variable, such as in the case of the visual ratings, and is particularly useful in studies where plot status is classified into categories that function like qualitative gradients, such as poor—moderate—good—exceptional [28].

2.6.2. Implementation of the Ordinal Regression

We conducted the fitting of our models using the R *ordinalForest* version 2.4-4 package [29]. The variable importance byproduct was extracted from each model that was fitted and used to rank predictors according to their overall impact on the model's performance. This R package can also conduct the prediction of ordinal response variables in statistical problems with few (low-dimensional) as well as many (high-dimensional) predictors in the dataset [29].

2.6.3. Description of Model Structures

We tested if using all available observations taken throughout the year would yield a strong predictive model. We did that by splitting the dataset into training (70% of observations) and test (the remaining 30%) sub-datasets. We named this model "Total". We also tested seasonal model variations. Each turf growing season was split into three sections: spring (April–June), summer (June–August), and fall (September–October). We

tested if observations collected in the Spring season as a training sub-dataset would predict Summer and Fall turf quality scores. In this case, the Summer and Fall seasons served as validation sub-datasets. We named this model "Spring". Likewise, the "Summer" and "Fall" models were tested. Finally, we were also interested in assessing if a simpler model would yield relatively acceptable predictions. We fitted a reduced version of the "Total" model but with a subset of the most important predictors as per our assessment of the top five predictors based on the rank probability scores (RPS) as described [29]. We named this model "FiveTotal".

2.7. Assessment of the Goodness of Fit

We organized the results from the different validations into confusion matrices (CM) to assess the goodness of fit [30] of each model. We utilized the R *caret* version 6.0-94 package [31,32] with the function *confusionMatrix* to extract the overall statistics, namely overall accuracy and the Kappa statistic for each model. To account for the drift in observations caused by different individual raters, we implemented a fuzzy accuracy assessment [33,34] on each one of the validation datasets. Any values that fell outside of the diagonal (i.e., improperly predicted visual rating values) were moved to the diagonal of the matrix (Figure 3) with customized R scripts. This was done, however, only if there was a positive or negative drift of one unit. For instance, if the originally observed value was 4 and the predicted value was 3 (negative drift) or 5, then the values were moved from the cell outside (i.e., prediction errors) the diagonal of the matrix to the cells in the diagonal (i.e., accurate predictions) of the matrix. However, if the originally observed value was 4 and the predicted value was 2 or 6, then those values remained in their cells and thus were included as prediction errors.

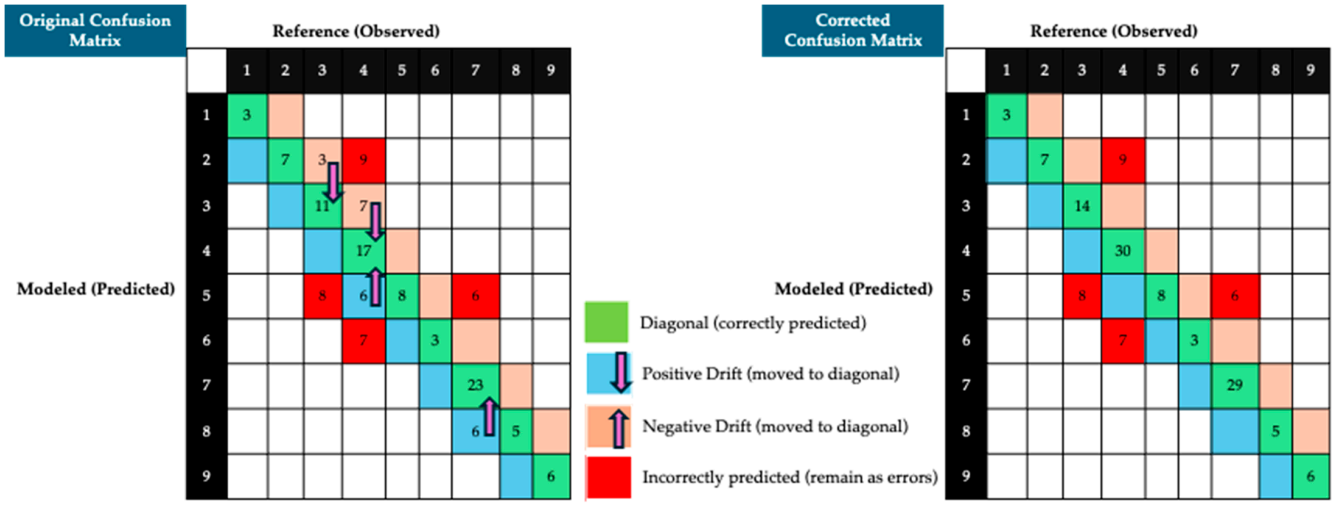

Figure 3. Schematic of how the fuzzy correction for the confusion matrices works. Predictions that are in one-value drift cells are moved to the diagonal and thus reduce the overall error, while predictions in two-value or more drift cells remain unchanged and keep contributing to the overall error.

Comparison Across Models

We computed the numeric value of the area under the receiver operating characteristic (ROC) or area under the curve (AUC) for each one of the five model structures that we tested. Given that this research problem dealt with nine ordinal classes, a regular two-classes AUC computation does not apply in this case; rather, a multi-class AUC approach [35] was utilized. A multi-class AUC is a mean of several two-classes AUC (i.e., the comparison of visual rating 1 with visual rating 2, or the comparison of visual rating 1 and visual rating 3, and so on). Given the nature of this multi-class AUC, it cannot be plotted, and only the area can be computed. We utilized the *multiclass.roc* function from the package *pROC* version 1.18.5 [36] to obtain multi-class AUC for each of the five model structures tested.

3. Results

3.1. Distribution of Visual Ratings per Season of Collection

As described (Section 2.2), we collected visual ratings during five seasons in the years 2022 and 2023. The lowest mean visual ratings were observed in the summer season (average values of 4.23 and 4.4), which was expected with the visual quality of the plots being degraded due to the summer high temperatures and corresponding heat stress on the plants. The highest ratings were recorded in the autumn of 2023. The distribution of the visual ratings per year/season is provided in the following figure (Figure 4). Interestingly, there were very few to no low visual ratings (i.e., ratings below 3) for the fall season of 2023.

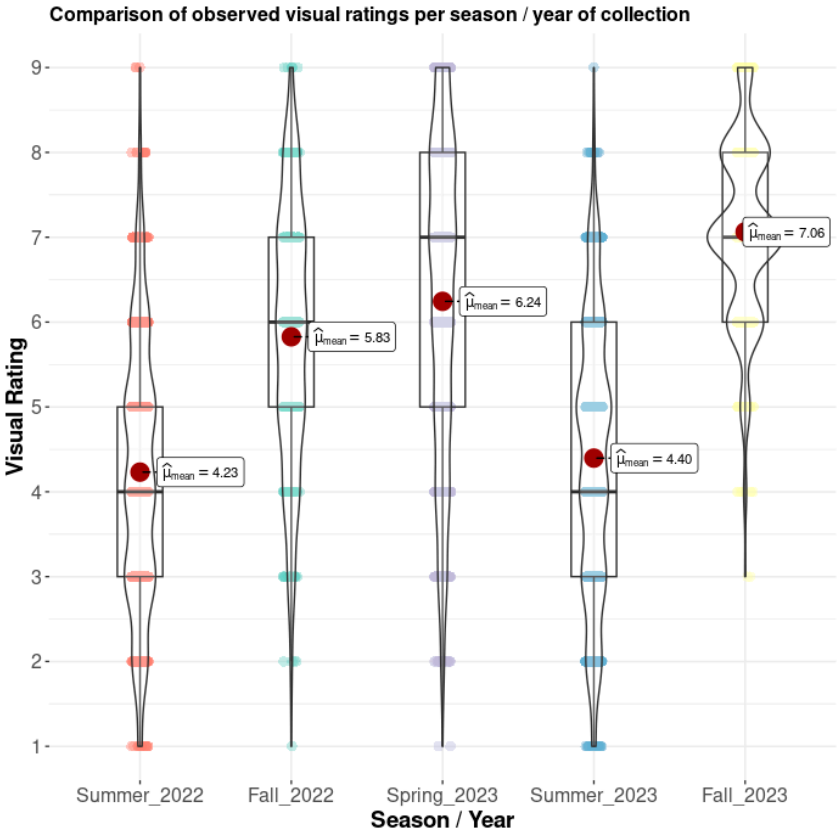

Figure 4. Temporal distribution of the visual ratings per season and year of collection.

More interesting are the seasonal transitions that have been observed for these plots during the observational period. We present a Sankey plot (Figure 5) that shows how the turf quality ratings have transitioned from one season to the next. For the sake of simplicity, we have collapsed the visual ratings into groups of three, denoting low quality (1, 2, 3), mid-quality (4, 5, 6) and high quality (7, 8, 9) ratings. The strata (vertical bars) indicate the proportion of each group of ratings per season, and the size of the flows between vertical bars indicate the proportion of each stratum that either (a) stayed in the same group or (b) transitioned to a different group. In general, there were no big transitions observed from high or mid-quality ratings to the low-quality group except from the spring of 2023 to the summer of 2023, which is quite understandable with all the stress that higher temperatures can bring about on the plants. In general, the transitions to the fall season represent a flow of recovery from the summer and stability within mid and high groups, as well as a flow from low to mid and from mid to high-quality ratings.

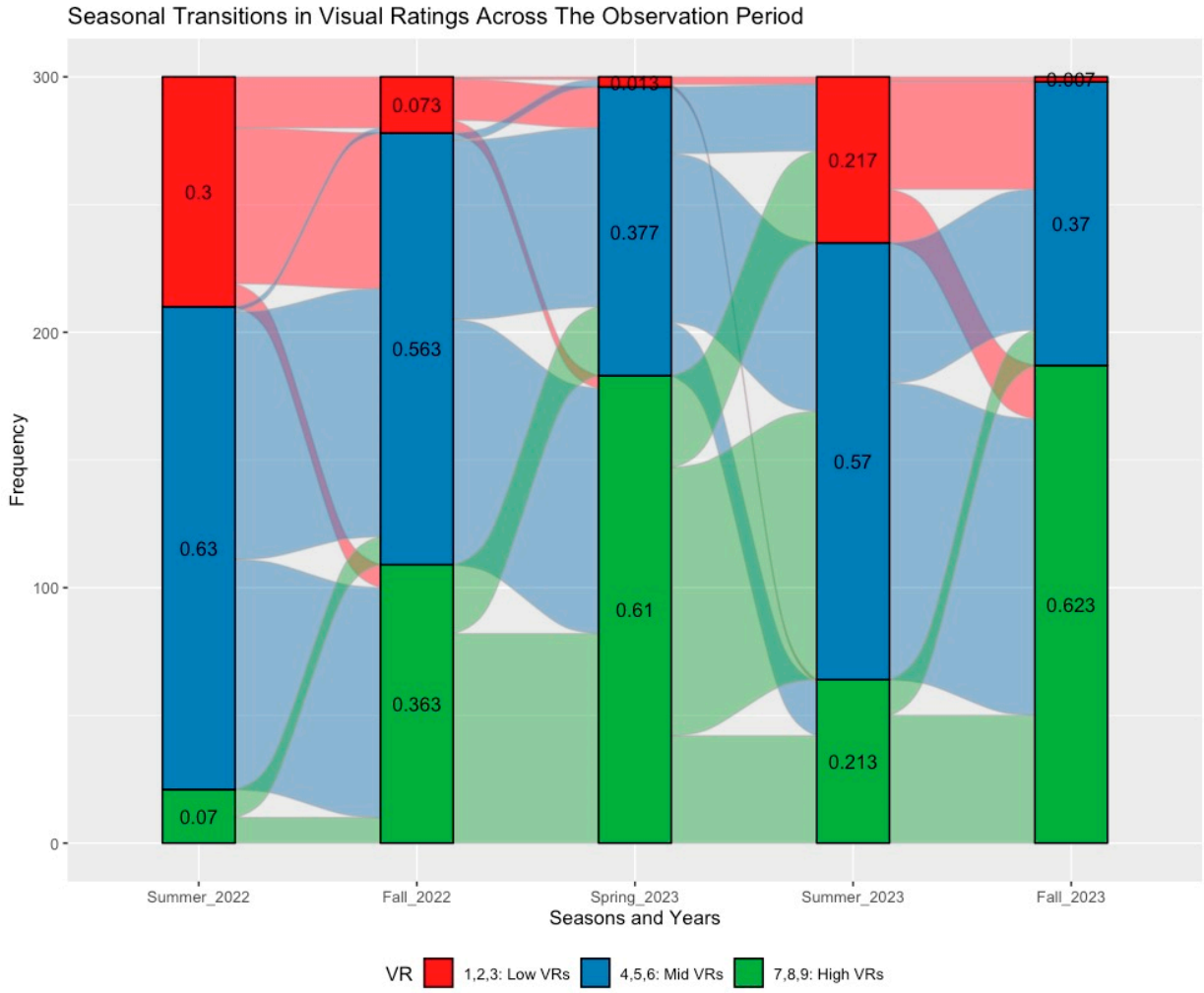

Figure 5. Temporal flows or transitions in turf quality visual ratings (VR) across seasons and years during the observational period.

3.2. Covariate Rankings—Variable Importance Across Models

For each one of the model variants that we fitted (Section 2.6.3), we extracted the rank probability scores (RPS). The RPS provides a reasonable indication of how important each predictor is within each model structure. Across all models, the green normalized difference vegetation index GNDVI was always among the five top predictors. Interestingly, the ratio vegetation index RVI, the normalized difference vegetation index NDVI and the chlorophyll index-green CIG, in addition to the GNDVI, were included in the top four predictors for both the Fall and the Total model structures (Figure 6). The thermal band shows as having significant importance in the "Summer" model (top five predictors) and in the top eight predictors for the "Fall" and "Total" model structures. This variable, however, showed a very low importance for the "Spring" model. Conversely, the dark green color index DGCI and the chlorophyll vegetation index CVI consistently ranked among the predictors with the lowest rank probability scores.

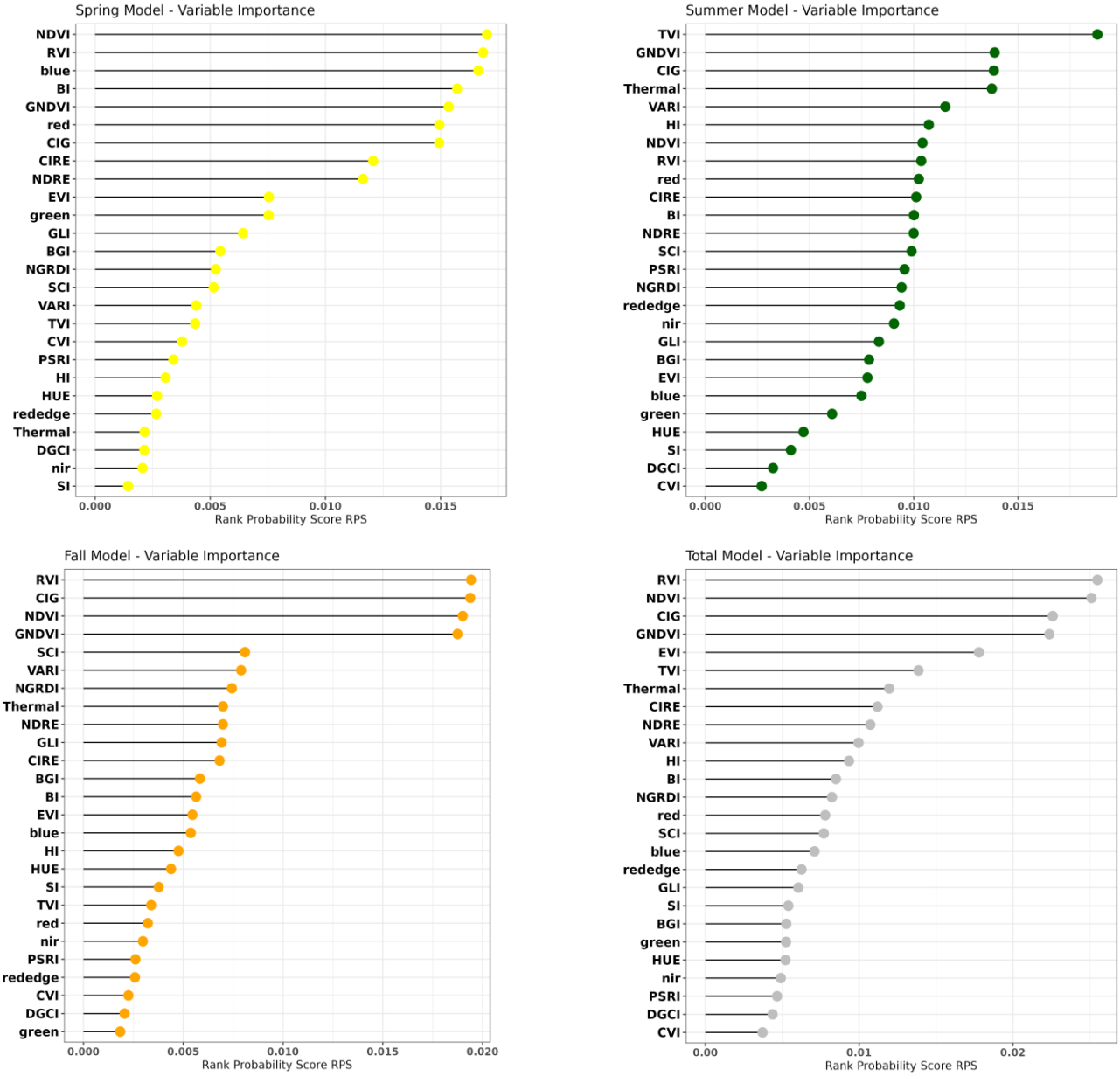

Figure 6. Variable importance plots for the model structures that were fitted in this research. Yellow = Spring Model, Green = Summer Model, Orange = Fall Model, Grey = Total Model.

3.3. Models' Performance—Uncorrected and Corrected Confusion Matrices

3.3.1. The "Spring" Model Structure

The overall accuracy for the Spring model was 0.207 for the uncorrected confusion matrix and 0.646 for the fuzzy-corrected confusion matrix. The kappa values were calculated as 0.0556 for the uncorrected matrix and 0.5758 for the fuzzy-corrected version. We also provide the following figure (Figure 7) that shows the original and the fuzzy corrected confusion matrices for this model structure. We include the per-class sensitivity values on the diagonal of each confusion matrix. It is clear that very poor per-class sensitivity values were obtained with the original (uncorrected) matrix. While the fuzzy-corrected matrix shows an improvement across all classes, it is noticeable that this model struggled with the prediction of the visual rating "4" (sensitivity of 8%), as well as with the visual rating "3" (sensitivity of 37%). The predictions of this model for the highest ratings (8 and 9) were moderate, with sensitivities in the 50% range.

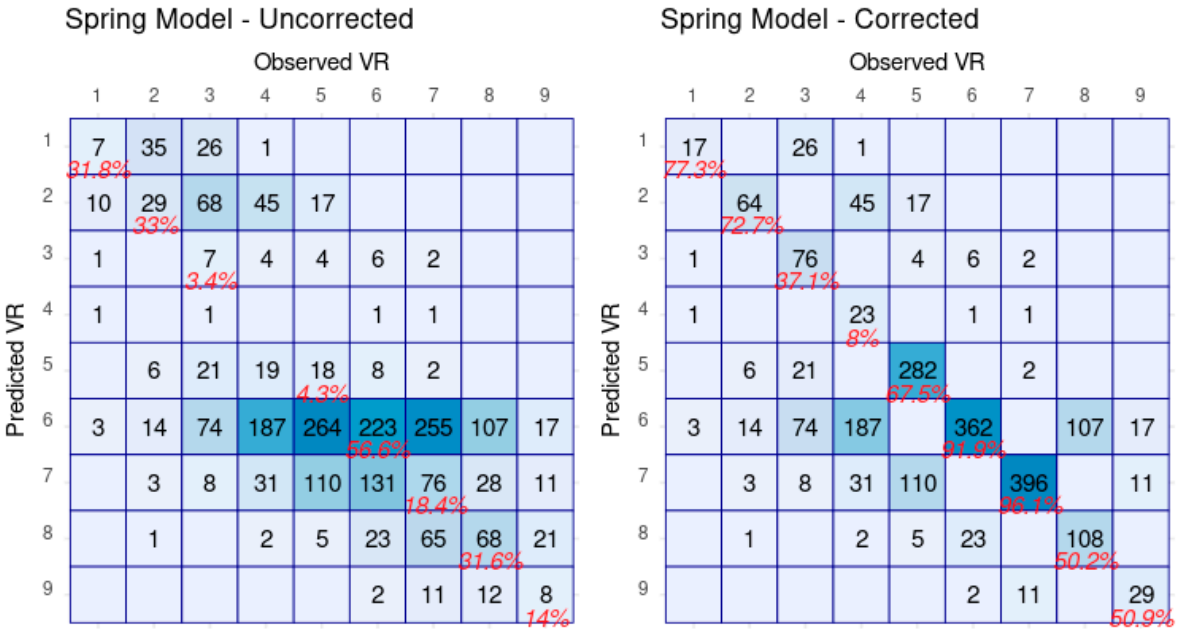

Figure 7. Uncorrected and corrected confusion matrices for the Spring model structure with per-class sensitivity values included.

3.3.2. The "Summer" Model Structure

This model exhibited the poorest results across all models. The overall accuracy for the "Summer" model was 0.189 for the uncorrected confusion matrix and 0.513 for the fuzzy-corrected confusion matrix. The kappa values were calculated as 0.0551 for the uncorrected matrix and 0.4311 for the fuzzy-corrected version. Similarly to the "Spring" model, the "Summer" model performed poorly with the middle (3, 4, 5) visual ratings, rendering sensitivity values lower than 40% (Figure 8).

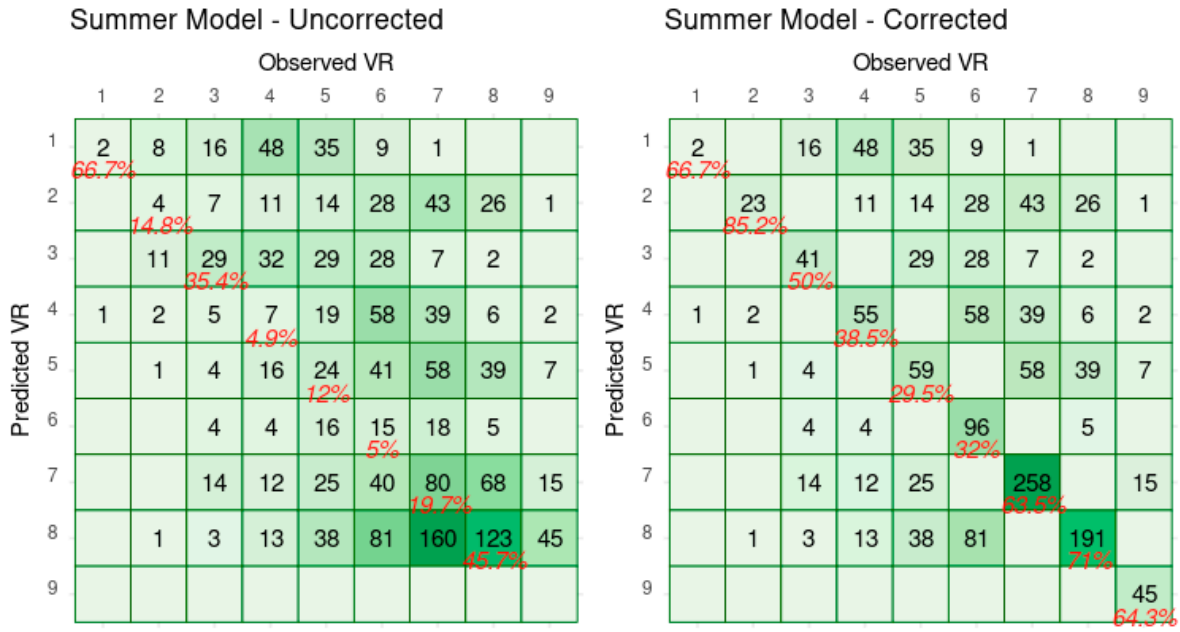

Figure 8. Uncorrected and corrected confusion matrices for the "Summer" model structure with per-class sensitivity values included.

3.3.3. The "Fall" Model Structure

This model exhibited the best results from the seasonal models. The overall accuracy was 0.287 for the uncorrected confusion matrix and 0.757 for the fuzzy-corrected confusion matrix. The kappa values were calculated as 0.1397 for the uncorrected matrix and 0.707 for the fuzzy-corrected version. While the model performed relatively well across all classes, it was not able to correctly discern the lowest visual rating: 1, with a sensitivity value of 4%. Further (Figure 9), and quite like the Spring model structure, this model achieved weak predictions for the upper visual ratings 8 and 9 with sensitivity values of 33% and 47%, respectively.

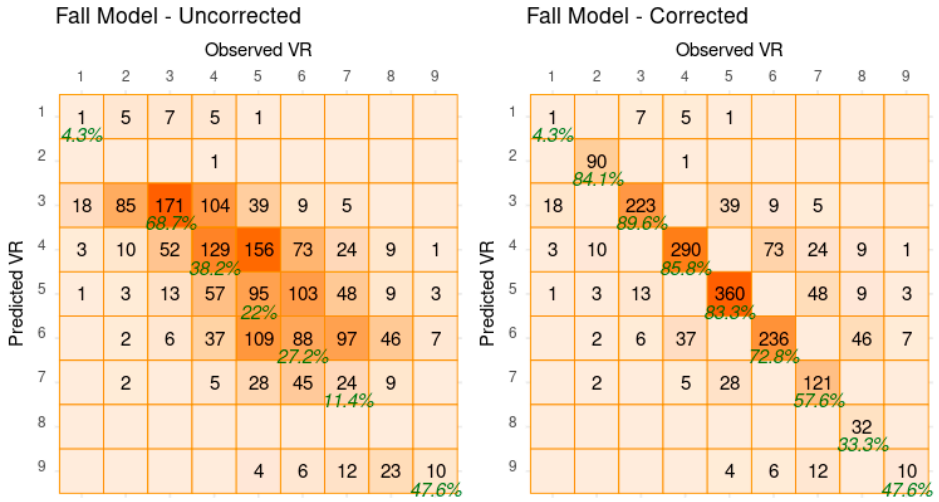

Figure 9. Uncorrected and corrected confusion matrices for the Fall model structure with per-class sensitivity values included.

3.3.4. The "Total" Model Structure

The best predictions were obtained when a model was built using visual rating observations drawn from the three seasons considered. This is the Total model structure. The overall accuracy was 0.463 for the uncorrected confusion matrix and 0.892 for the fuzzy-corrected confusion matrix. The kappa values were calculated as 0.3651 for the uncorrected matrix and 0.8726 for the fuzzy-corrected version. This model performed well (Figure 10) across all classes, with the lowest sensitivity value (71%) found for the lowest visual rating (1).

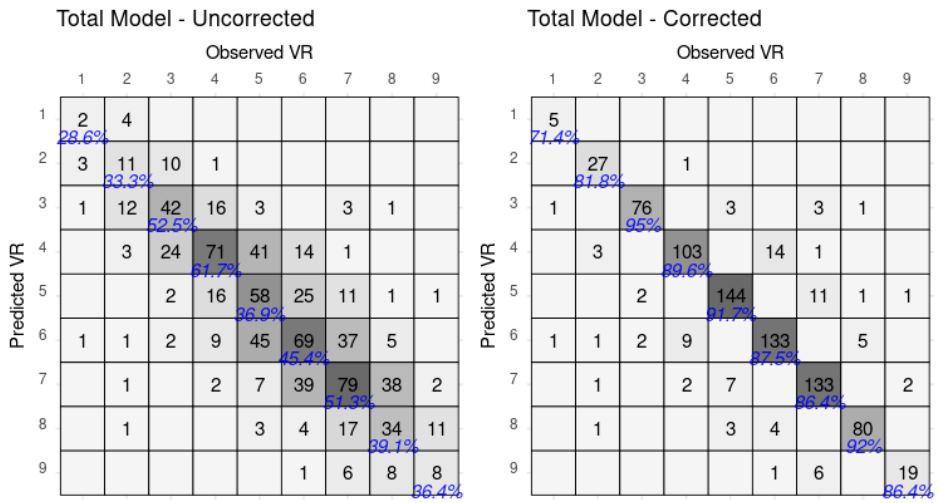

Figure 10. Uncorrected and corrected confusion matrices for the Total model structure with per-class sensitivity values included.

3.3.5. The "FiveTotal" Model Structure

Based on our assessment of the rank probability scores RPS (Section 3.2, Figure 4), we chose to keep the following predictors for the simpler model structure—vegetation indices: RVI, NDVI, GNDVI, CIG, and the thermal (LWIR) band. The overall accuracy was 0.42 for the uncorrected confusion matrix and 0.845 for the fuzzy-corrected confusion matrix. The kappa values were calculated as 0.3129 for the uncorrected matrix and 0.8167 for the fuzzy-corrected version. This simpler model structure performed (Figure 11) extremely well—comparable to the Total model that utilized the entire array of predictors, but with a much more parsimonious structure.

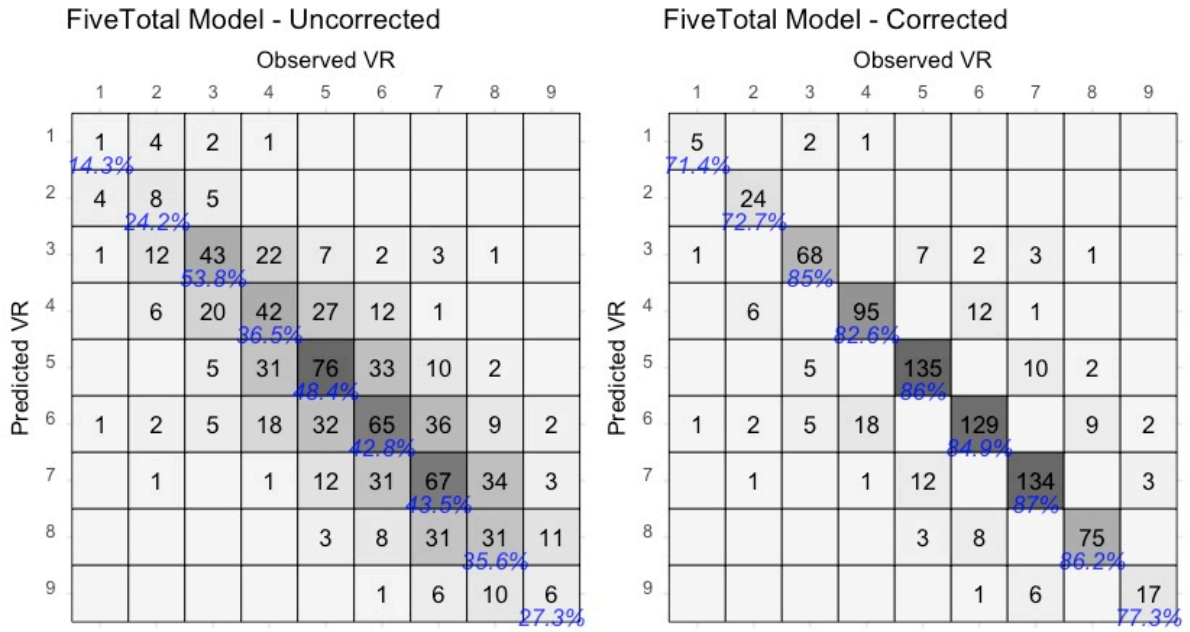

Figure 11. Uncorrected and corrected confusion matrices for the "FiveTotal" model structure with per-class sensitivity values included.

3.3.6. Inter-Model Comparison

We present the multi-class area under the curve AUC values for the different model structures in the following table (Table 3), along with the computed Kappa values previously reported. The multi-class AUC values confirm our previous statement that the "Summer" model structure rendered the worst predictions across all models and that the Fall model was the best model across seasonal structures. It is quite interesting that the reduced model (i.e., the Five-Total), which only used the top five predictors based on the rank probability scores, was highly comparable to the best overall model. The best model (i.e., Total) holds a positive difference of 0.016 AUC units over the Five-Total model, but the latter uses 21 fewer variables.

Table 3. Multi-class AUC (range 0.0–1.0) and Kappa values for the original and fuzzy-corrected confusion matrices for the model structures tested in this study.

Model	Multi-Class AUC	Kappa (Original)	Kappa (Fuzzy-Corrected)
Spring	0.759	0.056	0.576
Summer	0.673	0.055	0.431
Fall	0.774	0.139	0.707
Total	0.872	0.365	0.872
Five-Total	0.856	0.313	0.817

4. Discussion

4.1. On Selecting an Ordinal Regression Algorithm Instead of a Classification Approach

When presented with a research problem trying to fit a model for an ordinal or ranked variable, ordinal regression (OR) is a valued method with several advantages over traditional classification. A crucial strength of ordinal regression algorithms is the ability to identify significant predictor variables that influence the ordinal outcome, as OR allows for a detailed analysis of the direction and strength of these (i.e., predictors and response ordinal variable) relationships [37,38]. We were able to observe this advantage in our analysis. The chance to identify the five most important variables (Section 3.2) based on OR outputs (Figure 4) provided us with the ability to fit a reduced model (i.e., the Five-Total Model) that efficiently fitted the data and generated comparable AUC values to the Total model that used all the available predictors. Moreover, unlike standard classification methods that may overgeneralize the prediction assignment by regarding all ranks, ratings or classes as independent entities, OR considers the ordered nature of the outcome variable, which is indispensable for tasks where the relative ordering of classes carries valuable information [39]. This essentially applies to our research problem where the turf quality ratings are not just different from each other, but the rankings matter (i.e., the rating "3" is better than the rating "2" and inferior to the rating "4"). Further, OR acknowledges that the classes have a meaningful order and that the differences between adjacent classes are not necessarily equal. It has also been reported that the incorporation of this ordinal information into the modeling process delivers more accurate predictions and better captures the underlying structure of the data [39]. Modeling approaches that discard the ordinal information or assume equal spacing between adjacent ranks or classes may lead to poor performances when dealing with ordinal variables, which is not the case of OR that are designed to address the unique characteristics of ordinal data [40]. Furthermore, because OR explicitly accounts for the ordered nature of the classes, it can identify complex relationships and patterns in the data that traditional classification modeling approaches may ignore. This flexibility allows ordinal regression models to adapt to a wide range of ordinal prediction tasks and deliver more accurate and reliable results [41,42]. We considered the advantages of OR over classification methods in our selection of the modeling algorithm for our research analysis of turf quality visual ratings.

4.2. Seasonality and Its Impact on the Quality of Predictions from UAV-Multispectral Models

Our modeling performance results indicated that the model developed for the Fall season outperformed the other two seasons (Spring and Summer, Figures 7 and 8, respectively). Nevertheless, the "Fall" model only produced relatively acceptable results for the higher ratings (8 and 9), with an extremely poor performance for the lowest (1) visual rating. The fact that data used for the Spring model was collected only in one year while the Summer and Fall models' data were observed for two years contributed to the Spring model's accuracy being inferior with fewer observations and thus reduced temporal variability. The fall season is central for assessing turf quality, marking the transition period from summer recuperation to active growing to winter dormancy, thus making it a decisive time to evaluate the health and condition of the turf grasses. It has been suggested [14] that visual ratings during the fall provide valuable information into the overall performance (i.e., color, density, uniformity) of the turf as well as the responses of the grass to environmental stressors. Additionally, the fall season is a critical period for assessing the impact of environmental factors on turf quality, such as temperature, light availability, and nutrient availability [43]. Visual ratings collected during this time can help researchers understand how these factors influence turfgrass health and performance, providing valuable information for refining predictive models. The incorporation of visual ratings collected in the fall can improve the accuracy of predictive models as it allows the ability to integrate the response of turf grass to changing environmental conditions [43].

Based on the observed inter-seasonal transitions at our research site, the collection of visual ratings during the fall season becomes very important for understanding the

seasonal dynamics of turf quality, as they appear to give a baseline that can be used to compare ratings from other seasons. In our case, we were able to observe that the fall is a foundational season that acts as a bridge for recovery after the hot summer season and a path to either stability or improvement in overall quality to the spring after the dormancy in the winter. The type of predictive modeling that we propose in this paper seems to have benefited from this longitudinal (i.e., inter-seasons and years) data collection approach as it allowed robust predictions that account for seasonal variations in turfgrass performance. Still, in scenarios where field data collection must be limited to one season, our results seem to suggest that data from the fall can efficiently be used to model this ordinal variable. Based on our findings (Sections 3.3.3 and 4.1), by assessing turf quality during the fall, investigators can capture the seasonal dynamics of turfgrass performance.

4.3. Comparison Across Tested Model Structures and the Importance of Predictors

The best predictions (Table 3) for turf quality were obtained when a random sample was obtained to fit a model that included data across seasons and years (Section 3.3.4, Figure 10). This was the "Total" model. There was more data availability to be used for model fit, and this allowed the capture of the full seasonal variations (Figure 5) in both visual ratings and in the spectral and thermal signals obtained from the UAV imagery. Algorithms such as Random Forest or Ordinal Forest benefit from large amounts of samples available for model fitting. The best seasonal model was the "Fall" model structure, albeit this model rendered a poorer performance than the reduced "FiveTotal" model. This clearly suggests that to achieve the best accuracies in prediction, there is no need to separate the seasons. As discussed before (Section 4.2), the fall represents a bridge between turfgrass recovery and then dormancy, and this benefits the predictive modeling that was presented in this paper.

The "FiveTotal" structure (Section 3.3.5), which allowed to obtain the second-best prediction accuracies (Table 3), is based on five predictors: RVI, NDVI, GNDVI, CGI, and the thermal (LWIR) band, which have been documented to facilitate turf quality predictions. We found ample evidence of how these indices have been used in predictive modeling of turfgrass quality. Samples of research [9,44] have shown that the RVI and the NDVI can effectively differentiate between healthy and stressed turfgrass, making them reliable indicators for turf management practices. RVI has demonstrated a strong correlation with visual quality ratings, suggesting its utility in non-destructive assessments of plant health [45]. This correlation is particularly important for turfgrass species, where RVI has been used to monitor responses to environmental stressors like drought [46]. The GCI has emerged as a valuable tool for predicting turfgrass quality, primarily through its correlation with chlorophyll content, which is essential for photosynthesis and overall plant health. Studies [1] have documented that higher chlorophyll concentrations correlate with improved aesthetic quality, as chlorophyll is responsible for the vibrant green color of turfgrass, which is a key quality metric. The GNDVI has also been utilized in research [47] that has demonstrated strong correlations between UAV-derived GNDVI and ground-based assessments of turfgrass quality, with high R^2 values indicating reliable predictions. The use of thermal bands in UAV imagery has emerged as a significant method for assessing turfgrass quality. Thermal imaging allows for the detection of early drought stress in turfgrass, which is critical for effective management practices, particularly in environments like golf courses where maintaining quality is quite important. Research indicates [48] that thermal measurements can correlate with traditional visual assessments of turfgrass quality, providing a more objective means of evaluation. Furthermore, the integration of thermal imaging with other spectral data, such as NDVI, enhances the predictive capability of UAV-based imagery regarding turfgrass health and stress responses. While data for the spring was only collected in one year, we do not anticipate that this would impact the "FiveTotal" model as this model includes observations across seasons and years, and the only difference with the "Total" model is its utilization of a reduced, more parsimonious subset of predictors.

This multi-layered (thermal + spectral) approach not only aids in monitoring turf quality but also assists in optimizing turf management applications, thus promoting sustainable turf management practices. Overall, the combination of thermal and spectral data from UAVs presents a powerful tool for turfgrass managers aiming to maintain high-quality turf while minimizing resource inputs [47–49].

4.4. The Implications of Implementing Fuzzy Corrections on the Validation Datasets

Our implementation of a fuzzy correction (Section 2.7 Figure 3) to the resulting confusion matrices greatly impacted the original accuracies (i.e., overall, kappa) that were obtained from the different model structures. Proportionally speaking, the seasonal model structures (Spring, Summer and Fall) observed the highest increments in kappa (Table 3), where there were enhancements of 10-fold for the Spring model, 8-fold for the "Summer" model, and 5-fold for the Fall model structure. This was not the case with the Total and the Five-Total models, where the improvements in kappa were less than 3-fold, denoting that these model structures did not benefit as much as the seasonal models from the fuzzy corrections.

One significant advantage of using fuzzy corrections in confusion matrices is the ability to capture and deal with uncertainty and vagueness inherent in classification tasks or ordinal regression, as it was presented here. Traditional confusion matrices operate on precise and distinct classes, ranks or ratings. As we have explained throughout the text, this was not our case, as there can be inherent drift (positive or negative) in the visual ratings by different evaluators and across seasons. Operating on confusion matrices that use this assumption of unambiguity can lead to a loss of information regarding the degree of membership of instances to different classes [50]. By employing fuzzy logic, researchers can maintain this membership information, allowing for a more subtle understanding of classification or regression performance. It has been shown [50] that fuzzy confusion matrices could effectively address uncertainties in mapping grassland plant communities, thereby improving the accuracy of classification results. This capability of maintaining class membership information is particularly beneficial in remote sensing applications, where the boundaries between classes are often not well-defined [51], as was clearly our case.

In addition, it has been suggested [52] that the dynamic adjustment process of classes in the confusion matrices can lead to more reliable predictions, particularly where class distributions are imbalanced or when certain ranks or classes are challenging to predict, such as in the case of the Fall model with the higher ("8" and "9") quality ratings. However, the utilization of fuzzy corrections on confusion matrices is not without risks or challenges. One of those challenges is the reliance on expert knowledge about the magnitude of the corrections. Recall that in our fuzzy correction strategy (Section 2.7), we defined the positive or negative drift to be corrected to only one unit (i.e., just one level up or down of the diagonal of the matrix). This dependence on expert knowledge to define fuzzy membership functions can introduce subjectivity into the classification or regression process. This subjectivity can lead to inconsistencies in classification or ordinal regression results, as highlighted by the challenges faced in applying fuzzy logic in various domains [53]. Another potential disadvantage in the utilization of fuzzy corrections is the potential for ambiguity [54] in the interpretation of fuzzy-corrected confusion matrices. While the corrected matrices provide a relatively more transparent view of the performance of the ordinal regression, the nature of the results can make it challenging to draw clear-cut conclusions, as would be the case of the original non-corrected confusion matrices. This can impact the decision-making process and thus requires knowledge transfer to the final users for proper interpretation. We did not provide a thorough description of how to implement fuzzy corrections as it was out of the scope of this paper. However, comprehensive examples for the implementation of fuzzy corrections in the realm of land cover/land use classification can be found on [33] and also in [55].

4.5. The Opportunity to Complement Visual Ratings of Turf Quality with UAV Multispectral Imagery

While visual turfgrass quality ratings offer a quick and cost-effective method for evaluating turf quality, there are limitations that need to be considered. As mentioned earlier [9], there is the potential for intrinsic subjectivity as different evaluators may interpret and score turf quality differently, leading to inconsistencies in ratings. Further, relying solely on visual ratings may not deliver an ample evaluation of turf health as these visual assessments may not always detect issues such as nutrient deficiencies [10] or diseases. In situations like this, spectral and thermal information collected using UAVs may serve as complementary assessment methods to ensure an all-inclusive understanding of turfgrass conditions. UAVs can capture high-resolution multispectral images that provide a more objective measure of turfgrass quality, as this approach allows for the simultaneous assessment of various parameters such as color, density, and uniformity [47,56]. In addition, UAVs can cover relatively large tracks of land in a matter of minutes, where conducting visual assessments may be impractical for larger projects. Additionally, UAV imagery can facilitate the monitoring and comparison of turfgrass health or quality ratings through time. Our research covered field data collection through several seasons across two years, which required significant resources. The imagery collected in this period has shown strong predictive ability with visual assessments, indicating that UAVs can effectively capture subtle changes [49] in turf quality that may not be easily observable through visual ratings alone. Continuous monitoring is essential in turfgrass management in semiarid or arid lands to track improvements or deterioration due to high temperatures, drought or other factors. Visual ratings may not always capture the gradual changes effectively. Such a limitation of the visual ratings highlights the necessity of incorporating measurement tools such as UAVs and multispectral and thermal imagery to complement visual assessments for a more thorough evaluation of turf performance. Nevertheless, the digital signal from UAV imagery alone cannot provide meaningful information about turfgrass quality, and it completely depends on proper calibration that can only be obtained from field visual ratings.

4.6. Practical Applications and Limitations of the Developed Models in Turfgrass Management

Our research provides a foundation for the development of multiple models that can be trained in time (such as the inter-seasonal collection of data presented in this paper) or in space (where visual ratings are collected at different locations) to assemble a library of predictive models of turfgrass quality. Such a library can then be used at locations and/or times with no visual ratings, but with UAV imagery only, and in this case, predictions of turfgrass quality can be obtained faster and more efficiently. Further, the integration of machine learning models, such as our ordinal forests implementation with UAV data, may allow for dynamic adjustments in fertilizer applications [57] based on real-time turfgrass performance metrics, such as turfgrass quality. This can potentially optimize growth conditions and resource use. Similarly, the application of UAVs could hypothetically extend to irrigation management. It has been discussed [58] how UAV-based models can set standards to set up irrigation schedules. For instance, rapid turfgrass quality mapping and monitoring may pinpoint how turf is responding to different irrigation regimes. In this context, UAVs can facilitate the development of decision support tools that optimize water use while ensuring turfgrass quality is maintained. This application becomes relevant as harsher changes in the climate continue to impact water availability. Nevertheless, there are limitations in the utilization of the models that we presented here. The training and validation datasets were derived from a specific set of predictors in a single location over two years on a single turfgrass species and would have to be retrained to be applicable to other turfgrass quality evaluations. Research indicates that separate models may be necessary for different grass species to accurately predict turfgrass quality, as the relationships between UAV-derived data and ground measurements can vary significantly among species [47]. This species-specific variability greatly confounds the construction of a universal model that could be applicable across diverse turfgrass species. Another major

constraint is the complexity and expenses associated with implementing UAV technology in rapid turfgrass quality assessments. The initial investment in UAV systems, sensors, and data processing software can be relatively high. Further, the operation of UAVs requires specialized training that may not be available to all turfgrass managers and related stakeholders. These financial and technical obstacles can limit the accessibility of UAV technology, predominantly for smaller operations or where resources are limited [59]. In addition, while UAVs can collect imagery with exceptional spatial high-resolution, the frequency of data collection is severely limited by weather conditions (as explained in Section 2.3) and by restrictions dominated by airspace regulations [60]. These practical limitations can delay the ability to monitor turfgrass quality in near real-time, which is essential for managers as it drives timely interventions in turf management. Most of these limitations for implementing a UAV program were observed in this research project. As annotated earlier, we had contemplated starting the project in the spring of 2022. Nevertheless, our low level of technical proficiency in UAV-Sensor integration and the fact that UAV pilots were still being trained posed insuperable challenges for data collection during the spring season of 2022, and that is the main reason why data for this season was included only once in this research.

5. Conclusions

We have developed and validated seasonal as well as inter-seasonal predictive models of turf grass quality using ordinal regression for Kentucky Blue Grass. While the fall season model produced the best predictions from the individual seasonal models, it was clear that the inter-seasonal model that uses training data across seasons (spring, summer and fall) and years greatly outperforms the seasonal models, implying that more robust predictions can be obtained when seasonal and annual variations are included for model training. Further, our analysis of predictors' importance from the ordinal forest algorithm outputs pointed to the selection of five independent variables—four vegetation indices: NDVI, GNDVI, CGI, and RVI, and one of the original bands: the thermal band, which together in a model structure can rival the accuracy of the inter-seasonal model that utilizes the full array of predictors. We also implemented a fuzzy correction of the original confusion matrices to account for the positive and negative drift that may accompany the field-based turfgrass quality visual ratings. The fuzzy correction greatly impacted the original accuracies of the seasonal models (i.e., from 5-fold to 10-fold improvements), but this impact was not as significant for the inter-seasonal models, suggesting that even without these fuzzy corrections, the inter-seasonal (full and reduced) models generated relatively acceptable accuracies.

While visual ratings of turf quality are valuable in turfgrass management, they have limitations such as subjectivity, potential for bias, insensitivity to subtle changes, time-consuming nature, and variability in interpretation. On the other hand, the raw signals from UAV multispectral and thermal imagery cannot render turfgrass quality prediction by themselves, and they require quality data for training of predictive models. We have developed a methodological approach that can initiate the development of a predictive model library that can mature in the future to provide turfgrass predictions at times and places where only UAV imagery is available. This underscores the importance of complementing visual assessments with objective measurement techniques to ensure a more accurate and holistic evaluation of turf performance.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy14112575/s1. Table S1. List of vegetation indices used in this research.

Author Contributions: Conceptualization, A.H., P.J. and S.B.; methodology, A.H. and S.B.; software, A.H., K.P. and M.D.R.; validation, A.H.; formal analysis, A.H.; investigation, A.H. and S.B.; resources, A.H. and S.B.; data curation, A.H. and M.D.R.; writing—original draft preparation, A.H.; writing—review and editing, A.H., P.J., S.B. and M.D.R.; visualization, A.H.; supervision, A.H. and M.D.R.; project administration, S.B.; funding acquisition, S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: We would like to acknowledge the assistance of Matthew Kugel and Paul Harris, and the National Turfgrass Evaluation Program.

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

References

1. Dos Santos, P.L.F.; Silva, P.S.T.; Matos, A.M.S.; Alves, M.L.; do Nascimento, M.V.L.; de Castilho, R.M.M. Aesthetic and Sensory Quality of Emerald Grass (*Zoysia japonica*) as a Function of Substrate Cultivation and Mineral Fertilization. *Ornam. Hortic.* **2020**, *26*, 381–389. [CrossRef]
2. Bushman, B.S.; Robbins, M.D.; Thorsted, K.; Robins, J.G.; Warnke, S.E.; Martin, R.; Harris-Shultz, K. Transcript Responses to Drought in Kentucky Bluegrass (*Poa pratensis* L.) Germplasm Varying in Their Tolerance to Drought Stress. *Environ. Exp. Bot.* **2021**, *190*, 104571. [CrossRef]
3. Malinowski, D.P.; Belesky, D.P. Adaptations of Endophyte-Infected Cool-Season Grasses to Environmental Stresses: Mechanisms of Drought and Mineral Stress Tolerance. *Crop Sci.* **2000**, *40*, 923–940. [CrossRef]
4. Sandor, D.; Karcher, D.; Richardson, M.; Hignight, D.; Hignight, K. Kentucky Bluegrass Performance Under Chronic Drought Stress. *Crop Forage Turfgrass Manag.* **2019**, *5*, 180089. [CrossRef]
5. Braun, R.C.; Bremer, D.J.; Ebdon, J.S.; Fry, J.D.; Patton, A.J. Review of Cool-Season Turfgrass Water Use and Requirements: II. Responses to Drought Stress. *Crop Sci.* **2022**, *62*, 1685–1701. [CrossRef]
6. Wang, J.P.; Bughrara, S.S.; Nelson, C.J. Morpho-Physiological Responses of Several Fescue Grasses to Drought Stress. *HortScience* **2008**, *43*, 776–783. [CrossRef]
7. Morris, K.N.; Shearman, R.C. NTEP Turfgrass Evaluation Guidelines. In Proceedings of the NTEP Turfgrass Evaluation Workshop, Beltsville, MD, USA, 1998; pp. 1–5. Available online: https://www.ntep.org/pdf/ratings.pdf (accessed on 15 December 2022).
8. McNally, B.C.; Elmore, M.T.; Murphy, J.A.; Murphy, S.L. Annual Bluegrass and Creeping Bentgrass Tiller Response to Phosphate Fertilizer and Soil pH. *Crop Sci.* **2024**, *64*, 511–522. [CrossRef]
9. Leinauer, B.; VanLeeuwen, D.M.; Serena, M.; Schiavon, M.; Sevostianova, E. Digital Image Analysis and Spectral Reflectance to Determine Turfgrass Quality. *Agron. J.* **2014**, *106*, 1787–1794. [CrossRef]
10. Jespersen, D.; Leclerc, M.; Zhang, G.; Raymer, P. Drought Performance and Physiological Responses of Bermudagrass and Seashore Paspalum. *Crop Sci.* **2019**, *59*, 778–786. [CrossRef]
11. Karcher, D.E.; Richardson, M.D. Quantifying Turfgrass Color Using Digital Image Analysis. *Crop Sci.* **2003**, *43*, 943–951. [CrossRef]
12. Serena, M.; Schiavon, M.; Sallenave, R.; Leinauer, B. Nitrogen Fertilization of Warm-season Turfgrasses Irrigated with Saline Water from Varying Irrigation Systems. 1. Quality, Spring Green-up and Fall Colour Retention. *J. Agron. Crop Sci.* **2018**, *204*, 252–264. [CrossRef]
13. McCall, D.S.; Zhang, X.; Sullivan, D.G.; Askew, S.D.; Ervin, E.H. Enhanced Soil Moisture Assessment Using Narrowband Reflectance Vegetation Indices in Creeping Bentgrass. *Crop Sci.* **2017**, *57*, S-161–S-168. [CrossRef]
14. Taşkın, S.Z.; Bilgili, U. Effects of different nitrogen sources on turf quality and plants growth of some warm-season turfgrasses. *TJFC* **2022**, *27*, 167–174. [CrossRef]
15. Bertucci, M.B.; Karcher, D.E.; Richardson, M.D.; O'Brien, D.P. Evaluation of Newly Established Buffalograss for Tolerance to Glyphosate. *Crop Forage Turfgrass Manag.* **2022**, *8*, e20161. [CrossRef]
16. Woodget, A.S.; Austrums, R.; Maddock, I.P.; Habit, E. Drones and Digital Photogrammetry: From Classifications to Continuums for Monitoring River Habitat and Hydromorphology. *Wiley Interdiscip. Rev. Water* **2017**, *4*, e1222. [CrossRef]
17. Lee, H.; Bremer, D.J.; Su, K.; Keeley, S.J. Relationships between Normalized Difference Vegetation Index and Visual Quality in Turfgrasses: Effects of Mowing Height. *Crop Sci.* **2011**, *51*, 323–332. [CrossRef]
18. Fiorio, S.; Macolino, S.; Leinauer, B. Establishment and Performance of Bluegrass Species and Tall Fescue under Reduced-Input Maintenance in a Temperate Mediterranean Environment. *HortTechnology* **2012**, *22*, 810–816. [CrossRef]
19. Meza, K.; Torres-Rua, A.F.; Hipps, L.; Kustas, W.P.; Gao, R.; Christiansen, L.; Kopp, K.; Nieto, H.; Burchard-Levine, V.; Martín, M.P. Spatial Estimation of Actual Evapotranspiration over Irrigated Turfgrass Using sUAS Thermal and Multispectral Imagery and TSEB Model. *Irrig. Sci.* **2023**, 1–24. [CrossRef]

20. National Oceanic and Atmospheric Administration Summary of Monthly Normals—Logan, Utah—1991–2020. Available online: https://www.ncei.noaa.gov/access/services/data/v1?dataset=normals-monthly-1991-2020&stations=USC00425186&format=pdf&dataTypes=MLY-TMAX-NORMAL,MLY-TMIN-NORMAL,MLY-TAVG-NORMAL,MLY-PRCP-NORMAL,MLY-SNOW-NORMAL (accessed on 8 September 2024).
21. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An Efficient Alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.
22. Lowe, D.G. Object Recognition from Local Scale-Invariant Features. In Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, 20–27 September 1999; IEEE: Piscataway, NJ, USA, 1999; Volume 2, pp. 1150–1157.
23. ODM, A. OpenDroneMap 2020. Available online: https://github.com/OpenDroneMap/ODM (accessed on 30 May 2023).
24. Gillies, S. Rasterio: Geospatial Raster I/O for {Python} Programmers 2013. Available online: https://rasterio.readthedocs.io/en/stable/ (accessed on 21 October 2023).
25. Graesser, J.; Hardtke, L.; Mann, M.; Denham, R.; Xu, S. GeoWombat: Utilities for Geospatial Data 2021. Available online: https://geowombat.readthedocs.io/en/latest/install.html (accessed on 1 September 2023).
26. Baston, D.; ISciences, L.L.C.; Baston, M.D. Package 'Exactextractr'. *Terra* **2022**, *1*, 17.
27. McCullagh, P. Regression Models for Ordinal Data. *J. R. Stat. Soc. Ser. B (Methodol.)* **1980**, *42*, 109–127. [CrossRef]
28. Moe, S.J.; Couture, R.M.; Haande, S.; Solheim, A.L.; Jackson-Blake, L. Predicting Lake Quality for the next Generation: Impacts of Catchment Management and Climatic Factors in a Probabilistic Model Framework. *Water* **2019**, *11*, 1767. [CrossRef]
29. Hornung, R. Ordinal Forests. *J. Classif.* **2020**, *37*, 4–17. [CrossRef]
30. Congalton, R.G. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [CrossRef]
31. Kuhn, M. Building Predictive Models in R Using the Caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
32. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; Team, R.C. Package 'Caret'. *R J.* **2020**, *223*, 48.
33. Lowry, J.H.; Douglas Ramsey, R.; Stoner, L.L.; Kirby, J.; Schulz, K. An Ecological Framework for Evaluating Map Errors Using Fuzzy Sets. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 1509–1519. [CrossRef]
34. Gopal, S.; Woodcock, C. Theory and Methods for Accuracy Assessment of Thematic Maps Using Fuzzy Sets. *Photogramm. Eng. Remote Sens.* **1994**, *60*, 181–188. Available online: https://www.asprs.org/wp-content/uploads/pers/1994journal/feb/1994_feb_181-188.pdf (accessed on 1 July 2024).
35. Hand, D.J.; Till, R.J. A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems. *Mach. Learn.* **2001**, *45*, 171–186. [CrossRef]
36. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinform.* **2011**, *12*, 77. [CrossRef]
37. Leelawat, N.; Suppasri, A.; Charvet, I.; Imamura, F. Building Damage from the 2011 Great East Japan Tsunami: Quantitative Assessment of Influential Factors. *Nat. Hazards* **2014**, *73*, 449–471. [CrossRef]
38. Chen, C.-K.; John, H., Jr. *Using Ordinal Regression Model to Analyze Student Satisfaction Questionnaires. IR Applications, Volume 1, May 26, 2004*; Association for Institutional Research: Tallahassee, FL, USA, 2004.
39. Suárez, J.L.; García, S.; Herrera, F. Ordinal Regression with Explainable Distance Metric Learning Based on Ordered Sequences. *Mach. Learn.* **2021**, *110*, 2729–2762. [CrossRef]
40. Suzuki, H.; Gonzalez, O. Relative Predictive Performance of Treatments of Ordinal Outcome Variables across Machine Learning Algorithms and Class Distributions. *J. Behav. Data Sci.* **2022**, *2*, 73–98. [CrossRef]
41. Pérez-Ortiz, M.; Fernández-Delgado, M.; Cernadas, E.; Domínguez-Petit, R.; Gutiérrez, P.A.; Hervás-Martínez, C. On the Use of Nominal and Ordinal Classifiers for the Discrimination of States of Development in Fish Oocytes. *Neural. Process Lett.* **2016**, *44*, 555–570. [CrossRef]
42. Ruan, Y.-X.; Lin, H.-T.; Tsai, M.-F. Improving Ranking Performance with Cost-Sensitive Ordinal Classification via Regression. *Inf. Retr.* **2014**, *17*, 1–20. [CrossRef]
43. Chen, Z.; Wherley, B.; Reynolds, C.; Hejl, R.; Chang, B. Daily Light Integral Requirements for Bermudagrass and Zoysiagrass Cultivars: Effects of Season and Trinexapac-Ethyl. *Crop Sci.* **2021**, *61*, 2837–2847. [CrossRef]
44. Fitz–Rodríguez, E.; Choi, C.Y. Monitoring Turfgrass Quality Using Multispectral Radiometry. *Trans. ASAE* **2002**, *45*, 865. [CrossRef]
45. Wilber, A.L.; McCurdy, J.D.; Czarnecki, J.; Stewart, B.; Dong, H. Preemergence Herbicide Effects on St. Augustinegrass Establishment. *Agron. J.* **2023**, *115*, 1344–1355. [CrossRef]
46. Li, L.; Young, J.; Deb, S. Effects of Cultivation Practices and Products on Bermudagrass Fairways in a Semiarid Region. *Agron. J.* **2019**, *111*, 2899–2909. [CrossRef]
47. Zhang, J.; Virk, S.; Porter, W.; Kenworthy, K.; Sullivan, D.; Schwartz, B. Applications of Unmanned Aerial Vehicle Based Imagery in Turfgrass Field Trials. *Front. Plant Sci.* **2019**, *10*, 279. [CrossRef]
48. Hong, M.; Bremer, D.J.; van der Merwe, D. Thermal Imaging Detects Early Drought Stress in Turfgrass Utilizing Small Unmanned Aircraft Systems. *Agrosyst. Geosci. Environ.* **2019**, *2*, 190028. [CrossRef]

49. Caturegli, L.; Corniglia, M.; Gaetani, M.; Grossi, N.; Magni, S.; Migliazzi, M.; Angelini, L.; Mazzoncini, M.; Silvestri, N.; Fontanelli, M.; et al. Unmanned Aerial Vehicle to Estimate Nitrogen Status of Turfgrasses. *PLoS ONE* **2016**, *11*, e0158268. [CrossRef] [PubMed]
50. Rapinel, S.; Rossignol, N.; Hubert-Moy, L.; Bouzillé, J.-B.; Bonis, A. Mapping Grassland Plant Communities Using a Fuzzy Approach to Address Floristic and Spectral Uncertainty. *Appl. Veg. Sci.* **2018**, *21*, 678–693. [CrossRef]
51. Comber, A.; Fisher, P.; Brunsdon, C.; Khmag, A. Spatial Analysis of Remote Sensing Image Classification Accuracy. *Remote Sens. Environ.* **2012**, *127*, 237–246. [CrossRef]
52. Trajdos, P.; Kurzynski, M. Weighting Scheme for a Pairwise Multi-Label Classifier Based on the Fuzzy Confusion Matrix. *Pattern Recognit. Lett.* **2018**, *103*, 60–67. [CrossRef]
53. Jurišić, M.; Plaščak, I.; Antonić, O.; Radočaj, D. Suitability Calculation for Red Spicy Pepper Cultivation (*Capsicum annum* L.) Using Hybrid GIS-Based Multicriteria Analysis. *Agronomy* **2020**, *10*, 3. [CrossRef]
54. Dubois, D. Representation, Propagation, and Decision Issues in Risk Analysis Under Incomplete Probabilistic Information. *Risk Anal.* **2010**, *30*, 361–368. [CrossRef]
55. Hagen, A. Fuzzy Set Approach to Assessing Similarity of Categorical Maps. *Int. J. Geogr. Inf. Sci.* **2003**, *17*, 235–249. [CrossRef]
56. Rockstad, G.B.; Austin, R.E.; Gouveia, B.T.; Carbajal, E.M.; Milla-Lewis, S.R. Assessing Unmanned Aerial Vehicle-based Imagery for Breeding Applications in St. Augustinegrass under Drought and Non-drought Conditions. *Crop Sci.* **2024**, *64*, 496–510. [CrossRef]
57. Zhou, Q.; Soldat, D.J. Evaluating Decision Support Tools for Precision Nitrogen Management on Creeping Bentgrass Putting Greens. *Front. Plant Sci.* **2022**, *13*, 863211. [CrossRef]
58. Colmer, T.D.; Barton, L. A Review of Warm-Season Turfgrass Evapotranspiration, Responses to Deficit Irrigation, and Drought Resistance. *Crop Sci.* **2017**, *57*, S-98. [CrossRef]
59. Bremer, D.J.; Sullivan, D.G.; Vines, P.L.; McCall, D.; Zhang, J.; Hong, M. *Considerations with Using Unmanned Aircraft Systems in Turfgrass*; Kansas State University: Manhattan, KS, USA, 2023.
60. Friell, J.; Straw, C. Comparing Ground-Based and Aerial Data at Field Scale during Dry down on Golf Course Fairways. *Int. Turfgrass Soc. Res. J.* **2022**, *14*, 377–384. [CrossRef]