Short running title: Mapping pasture vegetation

Mapping the spatial distribution of botanical composition and herbage mass in pastures using hyperspectral imaging

Yumiko Suzuki[1], Hiroshi Okamoto[2], Makoto Takahashi[3], Takashi Kataoka[2] and Youichi Shibata[2]

1 Graduate School of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan

2 Research Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan

3 Field Science Center for Northern Biosphere, Hokkaido University, Sapporo, Hokkaido, Japan

Correspondence

Yumiko Suzuki, School of Veterinary Medicine, Kitasato University, Towada, Aomori 034-8628, Japan.

Email: suzuki@vmas.kitasato-u.ac.jp

Tel: +81(0)176234371

Fax: +81(0)176238703

Present address: Yumiko Suzuki, School of Veterinary Medicine, Kitasato University, Towada, Aomori 034-8628, Japan.

Part of this study was presented at the 65th Meeting of the Japanese Society of

Grassland Science (March 2009) and the 2009 CIGR International Symposium of the Australian Society for Engineering in Agriculture (September 2009).

Abstract

Quantifying the vegetation heterogeneity in pastures is important to understand the behavior and performance of grazing animals. In this study, we developed a technique for mapping spatial distribution of botanical composition (perennial ryegrass, white clover and other plants) and herbage mass of perennial ryegrass in a grazed pasture using a hyperspectral imaging system, which can glean more detailed spectral information than conventional multispectral imaging that uses standard band frequencies. Plant portions were extracted from hyperspectral image using normalized difference vegetation index threshold and classified into the three plant species or species groups using linear discriminant analysis models. The overall success rate of plant extraction and discrimination between plant species or species groups was 91.6 and 80.3%, respectively. Herbage mass of perennial ryegrass was estimated using partial least squares regression analysis, where the estimated values showed a coefficient of determination of 0.60 against the observed values. In consequence of the models, our results suggested that the difference of leaf cell structure and photosynthesis pigments influenced discrimination between plant species or species group and estimation of the herbage mass. The results demonstrate the potential of the hyperspectral imaging system for creating high spatial resolution maps of vegetation characteristics in pastures.

Keywords

Field map; herbage mass; image processing; plant species; spectral analysis.

Introduction

Quantity and quality of forage in pasture are not uniform because they are affected by plant species, herbage mass, plant height and density. The nutrient requirements of the grazing cattle are different by growing stage, and grazing behavior is selective. Therefore, grazing management such as grazing period, grazing time, grazing interval and stocking rate must be correctly executed in consideration of pasture characteristics (quality and quantity of forage, botanical composition) to supply ample nutrients for cattle. Consequently, pasture characteristics provide essential information for achieving optimum grazing management.

Several studies have investigated the distribution of pasture characteristics. Herbage mass distributions were observed using an electronic capacitance probe with suitable accuracy (Ogura and Hirata 2001; Hirata 2002), which however requires considerable labor and time to obtain the distribution on a large-scale field. Nowadays, remote sensing is regarded as a useful tool for surveying various large-scale fields, including grassland. Airborne and satellite remote sensing were used by estimating and mapping sward characteristics (Kawamura *et al*. 2005, 2011; Tsutsumi *et al*. 2005; Vescovo and Gianelle 2006; Cho *et al*. 2007). Although the spatial distribution in a large-scale field was efficiently acquired using these methods, detailed distributions could not be obtained due to low spatial resolution. On the other hand, ground-based remote sensing can acquire information with high spatial resolution, and the information for a large-scale field can be acquired by using a moving vehicle. In addition, the method can be applied for site-specific field management because it provides real-time measurements. Ground-based hyperspectral imaging (GHI), in particular, has attracted considerable

attention because more detailed spectral information can be gleaned from GHI than from conventional multispectral imaging systems that use standard band frequencies. This method has been applied to estimate the nitrogen content and chlorophyll index, detect yellow rust disease in winter wheat and discriminate crop and weed in soybean field (Inoue and Penuelas 2001; Moshou *et al*. 2006; Okamoto *et al*. 2007; Suzuki *et al*. 2008a, 2008b). The authors demonstrated that GHI is a useful technique for obtaining agricultural field information with suitable accuracy. However, the studies did not consider the spatial distribution in large-scale field because they processed only spectral data. Therefore, development of GHI system would be necessary to understand the distribution in consideration of spectral and spatial analysis. The long-term objective of the present study is to develop a GHI system for monitoring the spatial heterogeneity of pasture characteristics on a large-scale field. This paper describes a technique for mapping the spatial distribution of both plant species or species group and herbage mass (dry matter) of the main grass species using hyperspectral imaging.

Materials and methods [level 1 heading]

Study site [level 2 heading]

The study was conducted at a pasture (2.0 ha) in the Field Science Centre for the Northern Biosphere, Experimental Farm, Hokkaido University, Sapporo, Hokkaido, Japan (43°04′N, 141°20′E). The pasture, dominated by perennial ryegrass (*Lolium perenne* L.) with some cover of white clover (*Trifolium repens* L.), was grazed by lactating dairy cows and fertilized three times a year with annual rates of 150 kg N, 80 kg P and 100 kg K ha^{-1}. The mean annual temperature during the study periods was

9.4°C.

Hyperspectral imaging [level 2 heading]

Hyperspectral image were acquired between 09.00 and 11.30 hours JST (Japan standard time) from April to September in 2007, 2008 and 2009 using hyperspectral camera (ImSpector V10; Specim Ltd., Oulu, Finland; spectral wavelength range, 360–1010 nm; resolution, 10 nm). The camera is an integrated combination of imaging spectrograph and matrix detector, and each pixel in an image contains 54 spectral waveband values. This imaging device, which functions as a push broom scanner, obtains only one spatial line (480 pixels) at a time. Therefore, its direction must be moved in order to scan a target area (Okamoto *et al*. 2006a, Suzuki *et al*. 2008b).

The images were acquired by two types of capturing systems, which were used to acquire the field-scale image generating a map and the small-scale image analyzing the pixel spectra (Okamoto *et al*. 2006a). The field-scale images were acquired on August 26, 2008 and May 19, 2009. The capturing system for this image was composed of the imaging device, digital video recorder and vehicle (JK 14 HP; Iseki & Co., Ltd., Ehime, Japan) (Figure 1a). The camera was mounted on the front of the vehicle at a height of 2.25 m above the ground and at 90º (angle of depression), and the images were acquired while driving the vehicle (0.13 m s^{-1}). The pixel resolution was 4.34 mm (driving direction) × 2.33 mm (vertical direction against the driving direction). The small-scale images were acquired at approximately regular intervals during the experimental periods. The capturing system for this image was composed of the imaging device, digital video recorder, electrically driven pan head (KD; Mizar Optical Instruments, Tokyo, Japan) and tripod (Figure 1b). The imaging device was mounted on the pan head

at a height of 1.00 m above the ground, and the images were acquired by rotary-motion scanning of the electrically driven pan head (0.02 m s^{-1}). The pixel resolution of small-scale images is 1.16 mm (scanning direction) × 1.04 mm (vertical direction against the scanning direction). In both systems, the iris of the lens was manually adjusted according to the illumination, and the images were recorded on digital video tape (720 × 480 pixels) by the digital video recorder as video data files (format: DV; frame rate: 29.97). To identify plant species, RGB color images were acquired by digital video recorder and digital camera.

Interpretation of acquired data would require optical correction because illumination during image acquisition changes by sunshine or cloud. In this study, the data was revised by performing statistical methods.

Measurements of herbage mass [level 2 heading]

Vegetation at each small-scale image acquisition area was cut at a quadrat of 0.5 × 0.5 m. The samples were separated into perennial ryegrass, white clover, other plants and dead material, and dried in a heat-treating oven (PS-22; Tabai Mfg. Co., Ltd., Osaka, Japan) at 70°C for 24 h to obtain dry weights. A total of 187 points (12 in 2007, 53 in 2008 and 122 in 2009) were measured for hyperspectral image and herbage mass throughout this study.

Mapping of plant species and herbage mass [level 2 heading]

For mapping the plant species or species group and estimating the perennial ryegrass herbage mass, plants were first extracted from the hyperspectral images by distinguishing them from the bare ground and dead material. Next, the plants were

classified into species or species group, and the herbage mass of perennial ryegrass was estimated (Figure 2). These processes were performed by software that was developed using Microsoft Visual C#. The software is based on an object-oriented framework designed by Okamoto *et al.* (2006a, 2006b) for agricultural hyperspectral image analysis. Pixel spectral analysis for discrimination and estimation was performed using R software (R Development Core Team 2007).

Plant extraction [level 3 heading]

Plant areas were extracted prior to discrimination between plant species or species group (Figure 3). In plant extraction, the pixel was shown in white if it was classified as a plant. On the other hand, it was shown in black if it was classified as a non-plant (bare ground and dead material). The pixel spectra between plant and non-plants differed significantly in the red light range and near-infrared (NIR) light range, and two waveband values (678 and 760 nm) were especially different (Figure 4). Therefore, the normalized difference vegetation index (NDVI), which was calculated from these values, was employed as a threshold for the plant extraction model (Suzuki *et al*. 2008a).

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{red})}{(\mathrm{NIR} + \mathrm{red})}$$

where NIR is the value at a wavelength of 760 nm, and red is the value at 678 nm.

A total of 2000 pixel spectra were collected for the discrimination between plant and non-plants (600 from 2007, 600 from 2008 and 800 from 2009). The samples were divided in half based on sample range for modeling and validation datasets. A threshold

for the extraction was determined by the NDVI calculated from the modeling dataset, and the threshold was evaluated by the validation dataset.

Discrimination between plant species or species group [level 3 heading]

For discrimination between plant species or species group, the plant areas were classified into three category groups: perennial ryegrass, white clover and other plants (Figure 5). The discriminant models were created from the category groups, which are identified by visual observation, expressing the response variable and the 54 spectral waveband values expressing the explanatory variables. Ten pixels of spectra for each category group were sampled from each image (100 images for perennial ryegrass, 100 for white clover and 120 for other plants), that is, 1000 spectra of perennial ryegrass, 1000 spectra of white clover and 1200 spectra of other plants were collected, and they were partitioned into modeling and validation datasets. The validation dataset consisted of a single sample (10 spectra) chosen from among all samples, and the modeling dataset consisted of the remaining samples. The spectra were normalized because the total intensity level of each spectrum varied according to the illumination (Okamoto *et al*. 2006b). The normalized spectral value (N_i) for the waveband number i (i = 1–54) was calculated as:

$$N_i = \frac{S_i - S_{min}}{S_{mean} - S_{min}}$$

where S_i is the original spectral waveband value, and S_{min} and S_{mean} are the minimum and mean spectral values, respectively.

The normalized spectra were classified into three category groups using discriminant analysis. In this discrimination, perennial ryegrass was extracted from the plant pixels, and then the remaining pixels were classified into white clover and other plants. Therefore, two submodels (submodel 1 for extraction of perennial ryegrass, submodel 2 for discrimination between clover and other plants) were employed for graded discrimination between plant species or species group. The submodels were developed by linear discriminant analysis (LDA) using wavebands which are selected by stepwise selection with Wilks' lambda criterion, which is an index for testing statistical significance. LDA is a statistical method for discriminating between categories that finds a linear discriminant function for multi-category pattern classification. The validation process of submodel was repeated *N* (*N* is the number of samples) times with different combinations of datasets. The success rate of the model was calculated based on the number of correct predictions over the total number of tests.

LDA and stepwise selection were performed using the "lda" function (Ripley 1996; Venables and Ripley 2002) in the "MASS" package and the "greedy.wilks" function (Mardia *et al*. 1979) in the "klaR" package for R software.

Estimation of herbage mass [level 3 heading]

After the perennial ryegrass areas were identified, the herbage mass of perennial ryegrass was estimated (Figure 6). The estimation model for the herbage mass was created from the observed herbage masses expressing the response variable and the 54 spectral waveband values expressing the explanatory variables. The mean spectra calculated from 100 pixel spectra in each image (12 images from 2007, 53 from 2008 and 122 from 2009) were employed for estimation of herbage mass. The spectral

waveband values of perennial ryegrass were normalized, and the model was developed by the partial least squares regression analysis (PLSR) and validated by cross validation (leave-one-out method). PLSR is an extension of principal component regression analysis, and it is a statistical method combining partial least squares (PLS) and multiple regression analysis (MLR). PLS was performed to obtain the latent variables that were calculated from between the response variable and the explanatory variables, and MLR was performed to develop model from the herbage masses and the latent variables. PLSR was performed using the "plsr" function (Martens and Næs 1989) in the "pls" package for R software.

Results and discussion

Plant extraction

The NDVI of plant samples ranged from 0.29 to 0.91, and the NDVI of non-plant (bare ground and dead material) ranged from −0.23 to 0.37. The NDVI threshold for plant area extraction was determined to be 0.23. The success rate in the validation was 91.6%. Consequently, plant pixels could be extracted from the hyperspectral images with a high degree of accuracy. In addition, these thresholds were applied for the extraction of bare ground or dead material, and the rate of contribution to grazing livestock production could be understood.

Discrimination between plant species or species group

Plant spectra were classified into three groups (perennial ryegrass, white clover and other plants) by two submodels. Most of the selected wavelengths for submodel 1,

which was used for extracting perennial ryegrass, belonged to the wavelength range over 700 nm, i.e. NIR wavelengths. Most of the selected wavelengths for submodel 2, which was used for discriminating between white clover and other plants, belonged to the wavelength range 420–580 nm and 716–853 nm, i.e. NIR and the visible wavelengths. The success rate of each submodel was 78.2–93.6%, and the overall success rate was 80.3% (Table 1). Generally, reflectance in the NIR light region is affected primarily by the leaf cell structure, whereas reflectance in the visible light region is mostly determined by the photosynthesis pigments (Gates *et al.* 1965). Therefore, our results suggested that the difference of leaf cell structure and photosynthesis pigments influenced discrimination between perennial ryegrass, white clover and other plants. Although some pixels were misidentified, the classification algorithm demonstrated acceptable accuracy.

Estimation of herbage mass

The agreement of the estimated herbage mass by the PLSR model with the observed mass was acceptable, although some data sets deviated from the 1:1 line, with the coefficient of determination (R^2) of 0.60, the standard error of cross validation (SECV) of 28.71 g m^{-2} and the relative error (RE = SECV/mean of normalized spectrum) of 0.21 (Figure 7). The standardized partial regression coefficients at NIR light region (791, 726, 928 and 704 nm wavelengths) were highs (Figure 8). Because the reflectance in the NIR light region changes by the leaf cell structure (Gates *et al*. 1965), our results suggested that the leaf cell structure is related to the herbage mass.

Previous studies also estimated the herbage mass by using spectral data (Zhao *et al*. 2007; Kawamura *et al*. 2008; Biewer *et al*. 2009). The herbage mass was estimated

using a spectrometer (spectral wavelength range: 350–2500 nm) with R^2 of 0.39–0.93. Our result indicated estimation accuracy equal to or higher than previous studies. The difference between this study and previous studies was ability to estimate the herbage mass of a particular plant species. In this study, only the herbage mass of perennial ryegrass could be estimated by performing discrimination between species or species group.

Spatial distribution map

Plant species/species group map corresponded with RGB color image acquired by digital video camera (Figure 9a–b), and detailed variability of the herbage mass in the entire field was recognized by the herbage mass map (Figure 9c). The maps reflected the actual pasture characteristics, and the spatial heterogeneity of species/species group and the herbage mass could be understood from the maps. Consequently, it was demonstrated that GHI is a useful technique for estimating the pasture characteristics. However, it is necessary to measure detail spatial distribution in the actual field for evaluation of the maps and estimation accuracy of large scale.

Conclusions

This paper described a technique for spatial distribution mapping of both the plant species/species group and the perennial ryegrass herbage mass. In the first step of the mapping process, plants were extracted from the hyperspectral image by discriminating between the plants and non-plants (bare ground and dead material) using NDVI thresholding. Next, the plants were classified into species or species group using LDA

models. Finally, the perennial ryegrass herbage mass was estimated using PLSR. As a result of the validation for plant extraction, the overall success rate was 91.6%, and plant pixels were correctly extracted from the hyperspectral images. For the discrimination between plant species or species group, the overall success rate of the validation was 80.3%, and many pixels were correctly identified. The results from discrimination models indicated that the difference of leaf cell structure and photosynthesis pigments influenced discrimination. The validation of the estimation of herbage mass resulted with R^2 of 0.60, SECV of 28.71 g m^{-2} and RE of 0.21, and the result of model indicated that the leaf cell structure influenced the herbage mass estimation. The maps of the plant species or species group and herbage mass reflected the actual pasture characteristics, and the spatial heterogeneity of the pasture characteristics was recognized from the maps. It would be demonstrated that the hyperspectral imaging system developed is a useful technique for monitoring pastures. This system will be applied to grazing management by evaluating model reproducibility and conforming map accuracy on a larger scale.

Acknowledgements

This study was funded by a grant from the Ministry of Education, Culture, Sports, Science and Technology, Japan (No. 17208022). We would also like to thank the staff of the Experimental Farm, Field Science Centre for Northern Biosphere, Hokkaido University.

References

Biewer S, Fricke T, Wachendorf M (2009) Determination of dry matter yield from legume-grass swards by field spectroscopy. *Crop Sci* 49: 1927–1936.

Cho MA, Skidmore AK, Corsi F, Van Wieren SE, Sobhan I (2007) Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *Int J Remote Sens* 30: 499–515.

Gates M, Keegan HJ, Schleter C, Weidner R (1965) Spectral properties of plants. *Appl Opt* 4: 11–20.

Hirata M (2002) Herbage availability and utilisation in small-scale patches in a bahia grass (*Paspalum notatum*) pasture under cattle grazing. *Trop Grassl* 36: 13–23.

Inoue Y, Penuelas J (2001) An AOTF-based hyperspectral imaging system for field use in ecophysiological and agricultural applications. *J Remote Sens* 22: 3883–3888.

Kawamura K, Akiyama T, Yokota H, Tsutsumi M, Yasuda T, Watanabe O, Wang S (2005) Comparing MODIS vegetation indices with AVHRR NDVI for monitoring the forage quantity and quality in Inner Mongolia grassland, China. *Grassl Sci* 51: 33–40.

Kawamura K, Watanabe N, Sakanoue S, Inoue Y (2008) Estimating forage biomass and quality in a mixed sown pasture based on partial least squares regression with waveband selection. *Grassl Sci* 53: 131–145.

Kawamura K, Watanabe N, Sakanoue S, Lee H, Inoue Y (2011) Waveband selection using a phased regression with a bootstrap procedure for estimating legume content in a mixed sown pasture. *Grassl Sci* 57:81–93

Mardia KV, Kent JT, Bibby JM (1979) *Multivariate Analysis*. Academic Press, London.

Martens H, Næs T (1989) *Multivariate Calibration*. Wiley, New York.

Moshou D, Bravo C, Wahlen S, West J. (2006) Simultaneous identification of plant stresses and diseases in arable crops using proximal optical sensing and self-organizing map. *Precis Agric* 2: 149–164.

Ogura S. Hirata M (2001) Two-dimensional monitoring of spatial distribution of herbage mass in a bahiagrass (*Paspalum notatum* Flügge) pasture grazed with cattle. *Grassl Sci* 48: 317–325.

Okamoto H, Sakai K, Murata T, Kataoka T, Hata S (2006a) Object-oriented software framework developed for agricultural hyperspectral imaging analysis. *Agric Inform Res* 15: 103–112. (In Japanese with English abstract.)

Okamoto H, Sakai K, Murata T, Kataoka T, Hata S (2006b) Development of remote sensing software based on hyperspectral imaging framework. *Agric Inform Res* 15: 219–230. (In Japanese with English abstract.)

Okamoto H, Murata T, Kataoka T, Hata S (2007) Plant classification for weed detection using hyperspectral imaging with wavelet analysis. *Weed Biol Manag* 7: 31–37.

R Development Core Team (2007) *R*: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available from URL: http://www.R-project.org/ [cited 28 July 2007].

Ripley BD (1996) *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge.

Suzuki Y, Okamoto H, Kataoka T (2008a) Image segmentation between crop and weed using hyperspectral imaging for weed detection in soybean field. *Environ Control Biol* 46: 168–174.

Suzuki Y, Tanaka K, Kato W, Okamoto H, Kataoka T, Shimada H, Sugiura T, Shima E (2008b) Field mapping of chemical composition of forage using hyperspectral

imaging in a grass meadow. *Grassl Sci* 54: 179–188.

Tsutsumi M, Kawamura K, Sugiyama M, Sato S, Deguchi Y, Sugawara K, Sakanoue S, Itano S (2005) Estimating the spatial distribution of available biomass in grazing forests with a satellite image: A preliminary study. *Grassl Sci* 51: 107–111.

Venables WN, Ripley BD (2002) *Modern Applied Statistics with S*, Edition 4, Springer, New York.

Vescovo L, Gianelle D (2006) Mapping the green herbage ratio of grasslands using both aerial and satellite-derived spectral reflectance. *Agr Ecosyst Environ* 115: 141–149.

Zhao D, Starks PJ, Brown MA, Phillips WA, Coleman SW (2007) Assessment of forage biomass and quality parameters of bermudagrass using proximal sensing of pasture canopy reflectance. *Grassl Sci* 53: 39–49.

Table 1 Validation results of plant species or species group classification algorithm

Model	Selected wavelength (nm)	Success rate (%)	Overall success rates (%)
Submodel 1	741, 728, 841, 778, 753, 866	93.6	80.3
Submodel 2	753, 420, 716, 433, 580, 543, 766, 853, 568	78.2	

Figure legends

Figure 1 Hyperspectral capturing system.

Figure 2 Flow of discrimination between plant species or species group and estimation of perennial ryegrass herbage mass.

Figure 3 Flow of plant area extraction.

Figure 4 Pixel spectra of plant, bare ground and dead material.

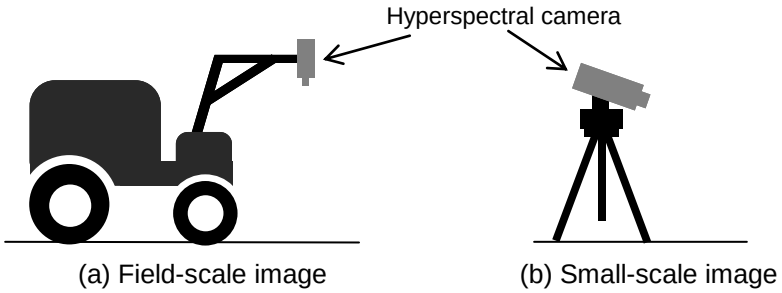
Figure 5 Flow of discrimination between plant species or species group.

Figure 6 Flow of estimating herbage mass of perennial ryegrass.

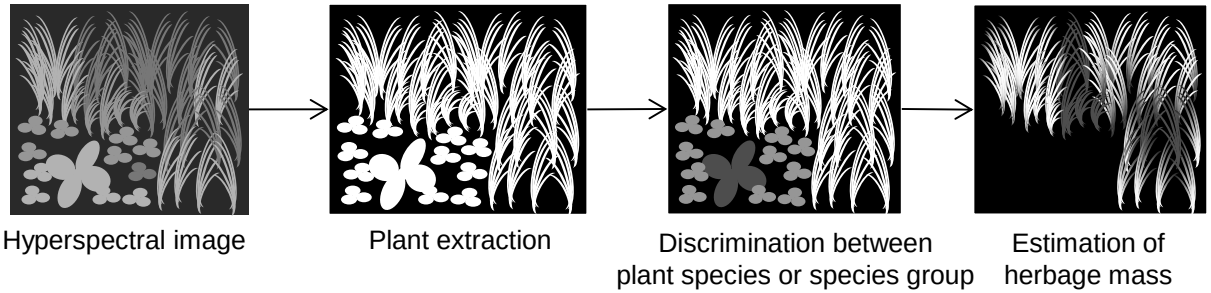
Figure 7 Relationship between the observed and estimated herbage masses in PLSR. The line shows the 1:1 relationship.

Figure 8 Standardized partial regression coefficients of 1st and 2nd latent variables retained in PLSR model.

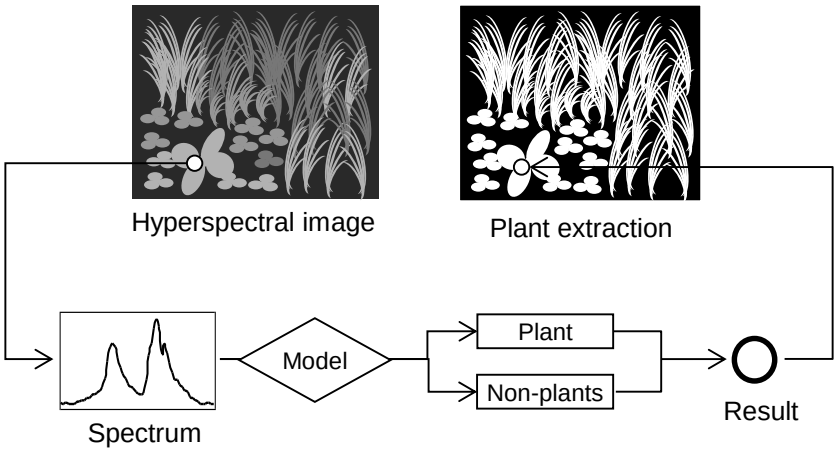
Figure 9 RGB color image, plant species or species group map and herbage mass map.



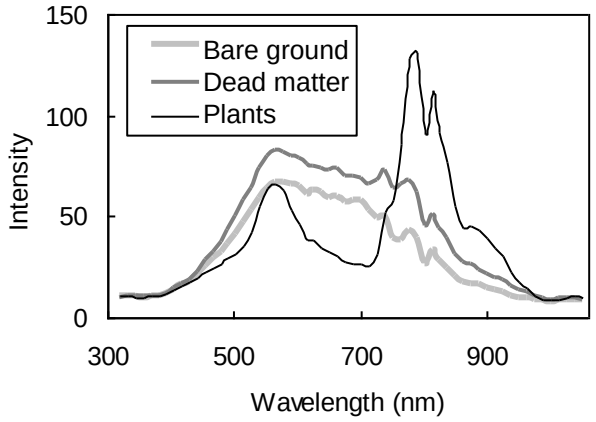
Suzuki et al. Figure 1



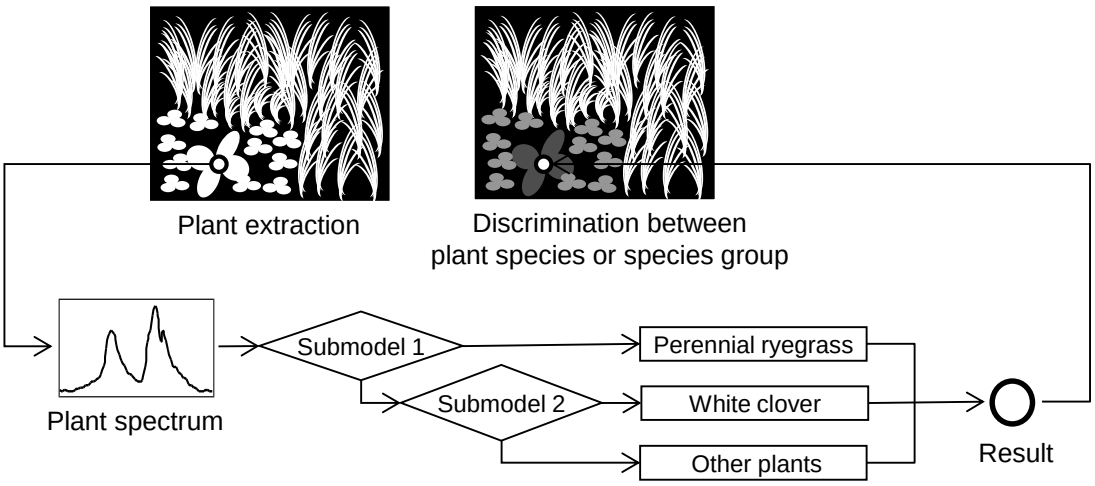
Suzuki et al. Figure 2



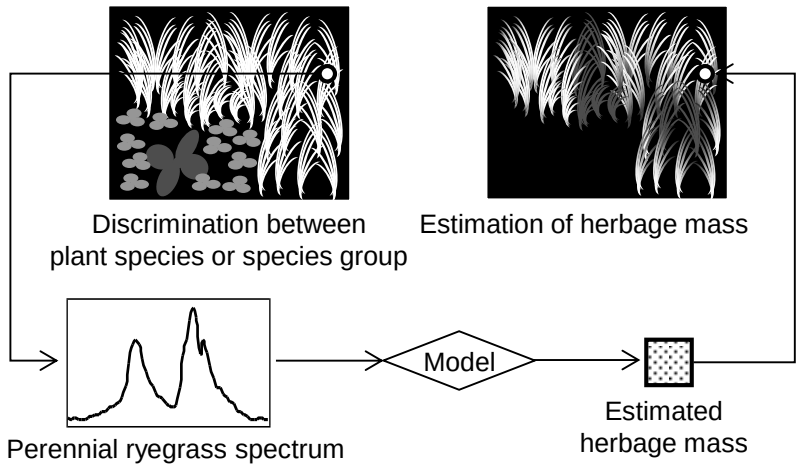
Suzuki et al. Figure 3



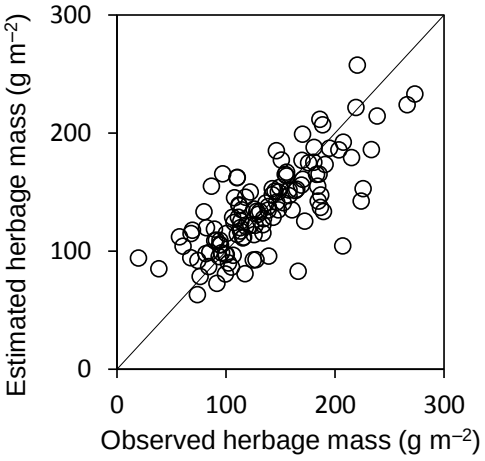
Suzuki et al. Figure 4



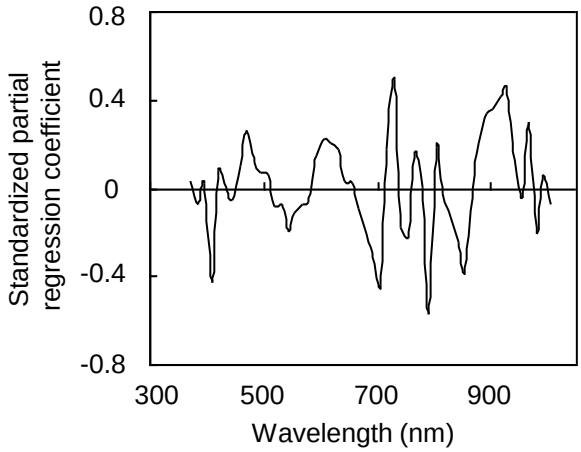
Suzuki et al. Figure 5



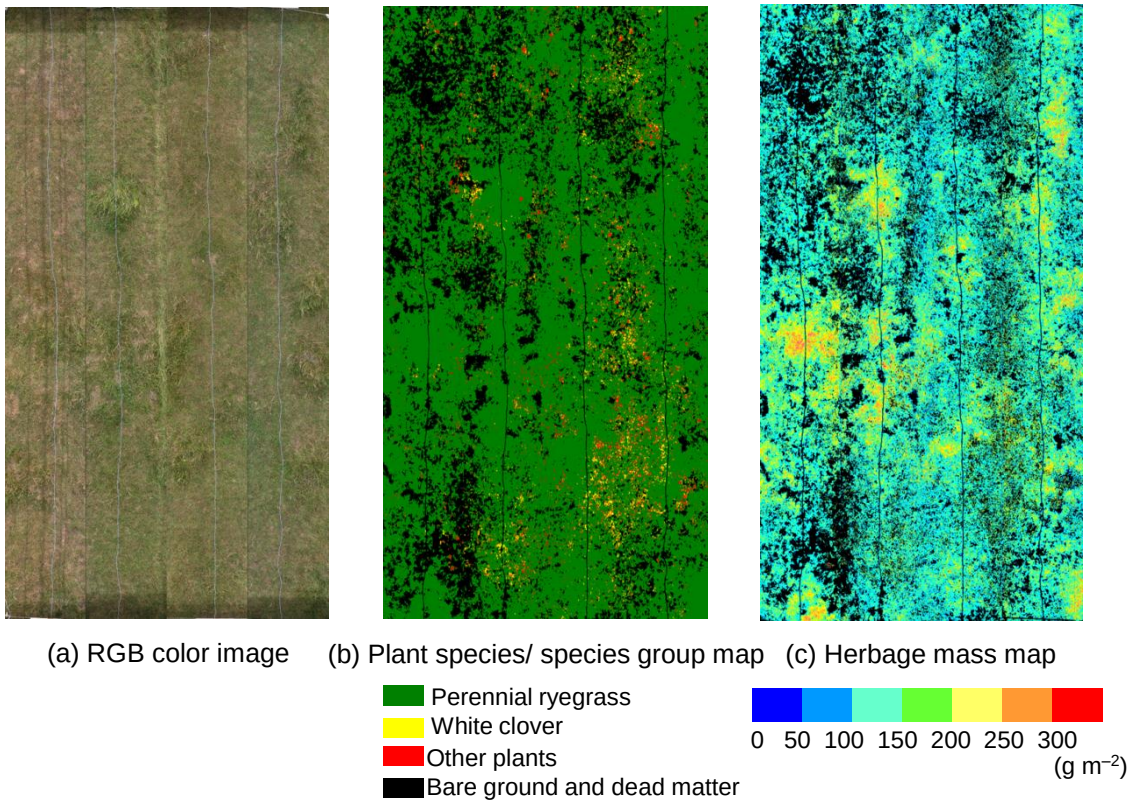
Suzuki et al. Figure 6



Suzuki et al. Figure 7



Suzuki et al. Figure 8



Suzuki at al. Figure 9