Is more data better? A comparison of multi- and hyperspectral imaging in phenotyping.

Marja Haagsma[1], Gerald Page[1], Jeremy Johnson[2,3], Chris Still[1], Kristen Waring[3], Richard Sniezko[2], John Selker[1]

1. Oregon State University 2. USDA Forest Service, Dorena Genetic Resource Center, 3. Northern Arizona University

Contact: haagsmam@oregonstate.edu

Introduction

Traditional methods of screening trees for resistance or susceptibility to pathogens have relied on many thousands of manual visual inspections of trees and seedlings. However, recent developments in hyperspectral (HS) remote sensing are showing great promise for the automated measurement of plant structural and physiological traits at a range of scales. HS sensors are still extremely expensive, while sensors with a lower spectral resolution and a lower cost price have also been successful in phenotyping applications. The objective of this study is to get a sense of the difference between using typical hyperspectral and multispectral data for a genotyping application. In this case study we aim to **I) detect infection** and **II) symptom severity** of white pine blister rust (caused by *Cronartium ribicola)* in southwestern white pine seedlings (*Pinus strobiformis*).

Method

Half of the seedlings (87) were inoculated, and the other half (88) was kept as non-inoculated control. Hyperspectral imaging acquisition started 9 months after inoculation and was repeated 16 times over the next 8 months. Seedlings were grown from 10 seed sources, resulting in distinct genotypes. This introduced a variety of resistance to blister rust.

Data

Hyperspectral data was captured with a VNIR Nano-Hyperspec camera (Headwall Photonics, MA, USA). Reflectance was measured in 271 bands, with 2.2 nm band center dispersion, 5-6 nm FWHM, and in the 400 to 1000 nm range. (Cost: ~$100,000)
Multispectral data was derived from the hyperspectral data set. This artificial data set has the same band center and width as the Micasense Rededge-MX sensor. It has 5 band in the 400 to 900 nm range with band widths ranging from 10 to 40 nm. (Cost: ~$5,000)
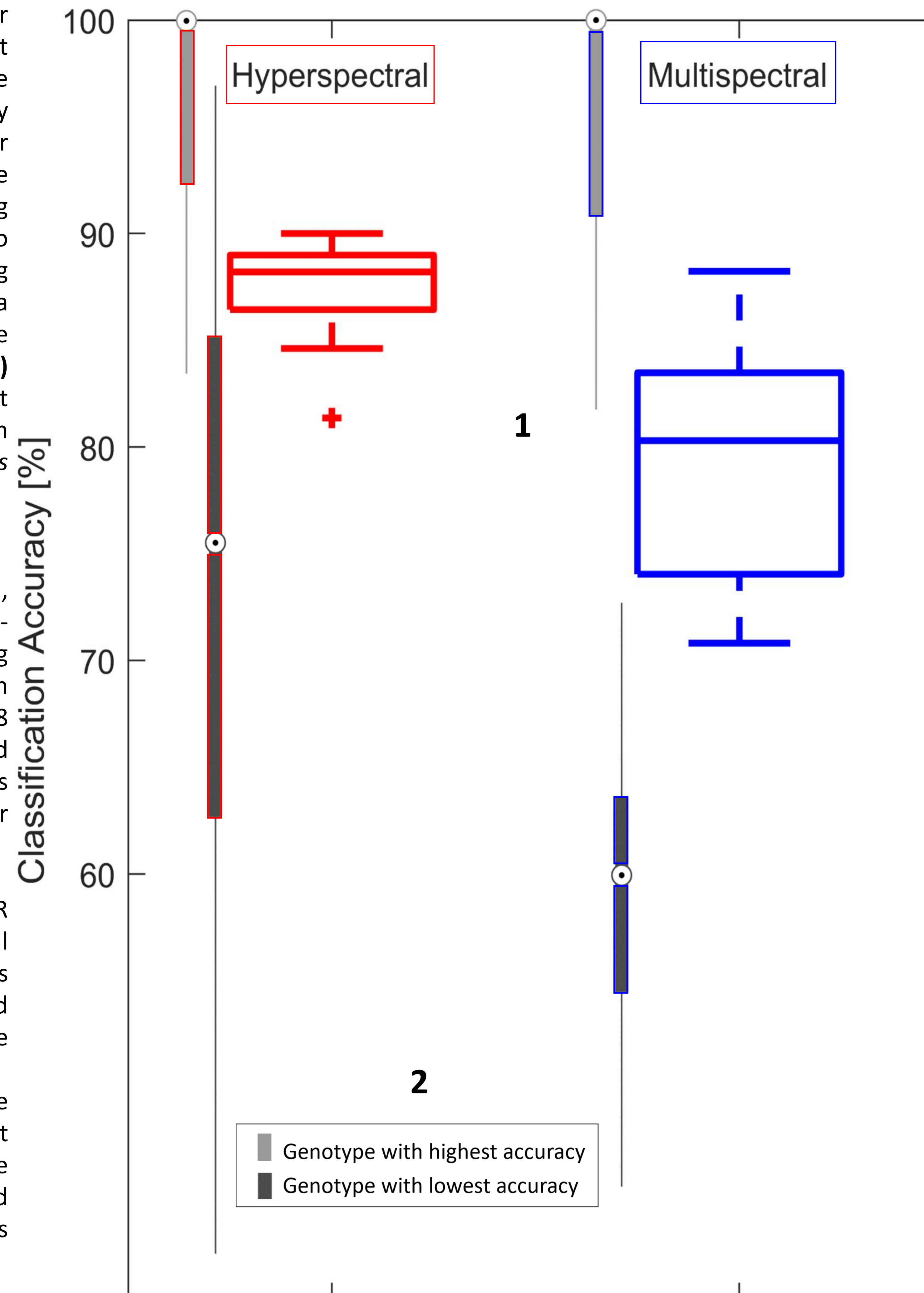
Analysis

Data was collected in array format. Per seedling, the median reflectance per band was extracted and used for classification. Per date the seedlings were classified for I) infection and II) symptom severity using a 'support vector machine' algorithm in Matlab (Mathworks v.2018b). The performance of the model was assessed using a 10-fold cross-validation. This was repeated 19 times and the final label was determined using the majority vote principle. Classification accuracy is the percentage of correctly predicted seedlings.
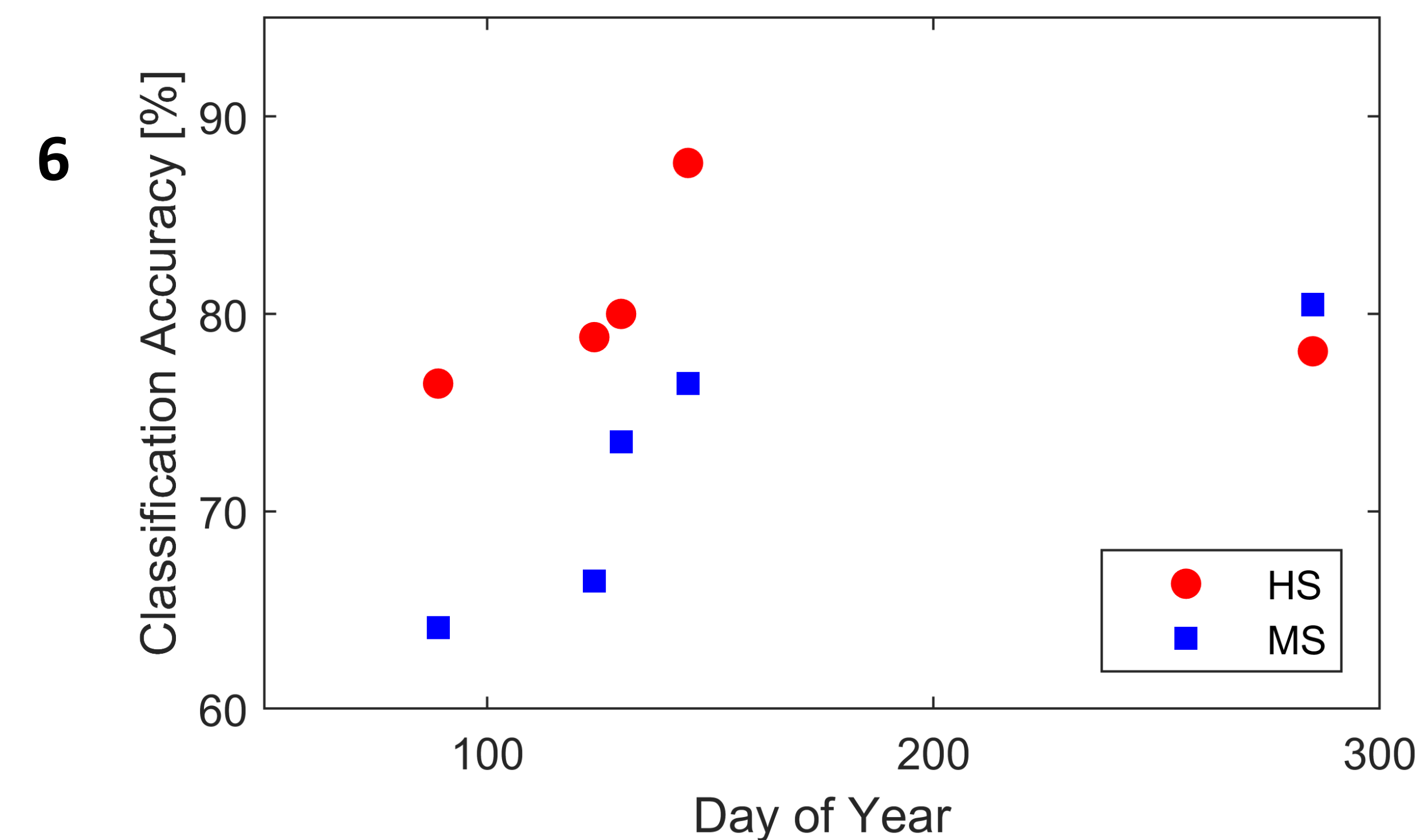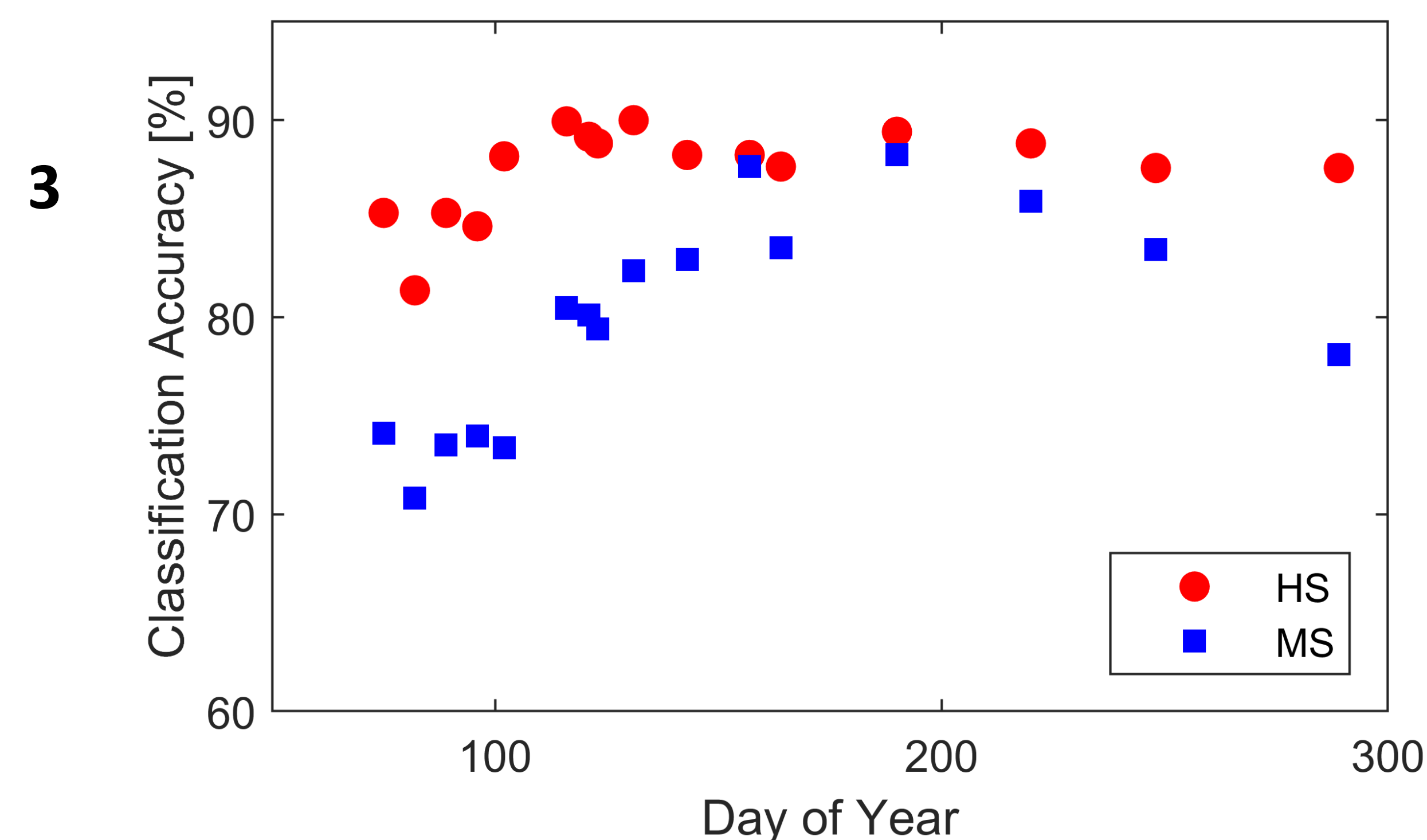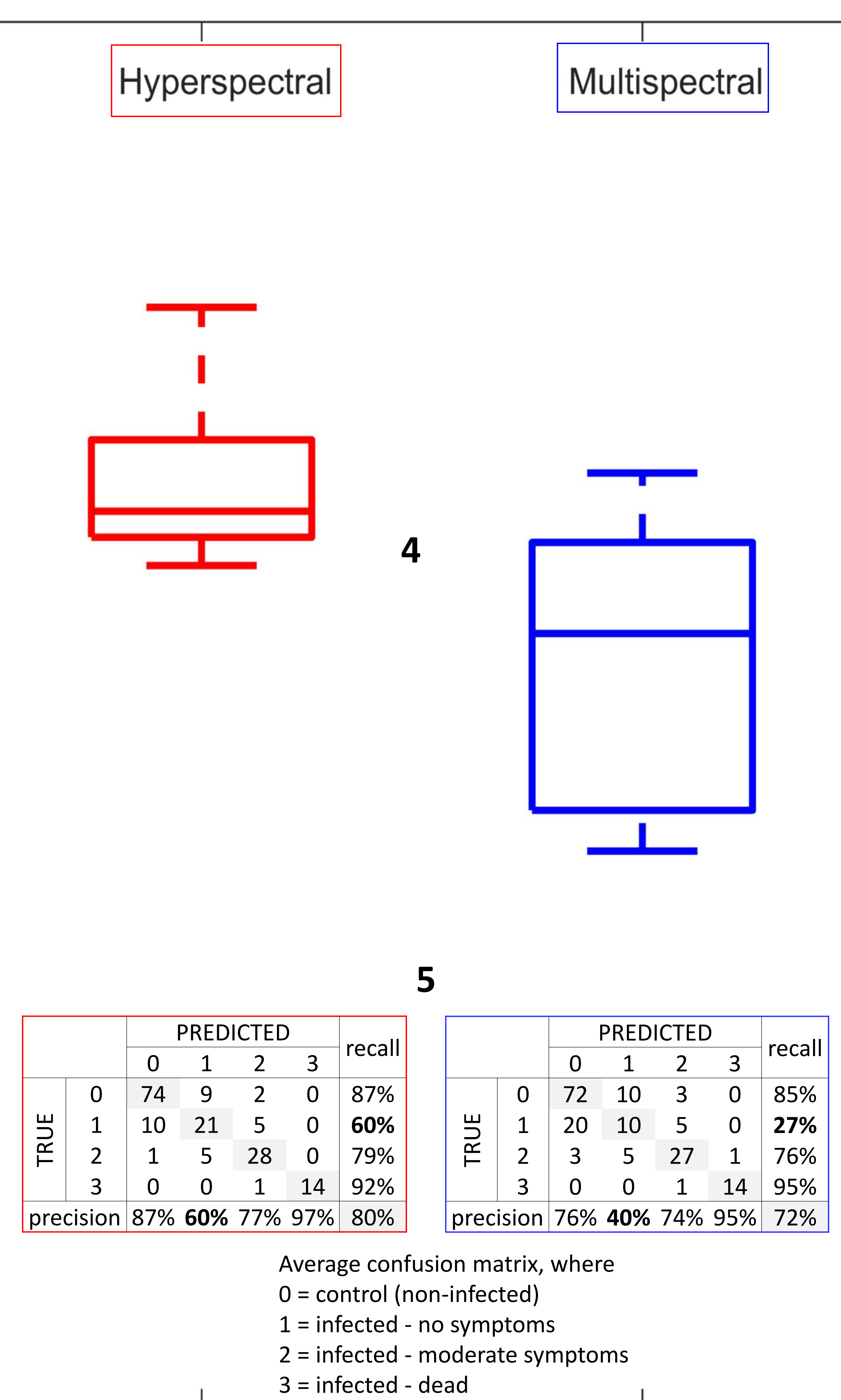
I. Infection detection

Hyperspectral / Multispectral

Classification Accuracy [%]

Genotype with highest accuracy
Genotype with lowest accuracy

II. Symptom severity

Hyperspectral / Multispectral

		PREDICTED				recall
		0	1	2	3	
TRUE	0	74	9	2	0	87%
	1	10	21	5	0	**60%**
	2	1	5	28	0	79%
	3	0	0	1	14	92%
precision		87%	**60%**	77%	97%	80%

		PREDICTED				recall
		0	1	2	3	
TRUE	0	72	10	3	0	85%
	1	20	10	5	0	**27%**
	2	3	5	27	1	76%
	3	0	0	1	14	95%
precision		76%	**40%**	74%	95%	72%

Average confusion matrix, where
0 = control (non-infected)
1 = infected - no symptoms
2 = infected - moderate symptoms
3 = infected - dead

I. Results

1. On average, the classification accuracy with HS imaging is 8 percent points higher than with MS imaging.
2. Even though the accuracy distribution is narrow in 1, a subset of two extreme genotypes (with the highest and lowest accuracy) has a wide range of accuracy. There is no difference in accuracy between HS and MS for the genotype with the highest accuracy. However, for the genotype with the lowest accuracy there was a significant reduction of accuracy when using MS over HS. This subset had a low number of individuals with symptoms.
3. At the beginning and end of the experiment period, classification accuracy using HS imaging is substantially higher than for MS data. This disparity is thought to be caused by seedlings that hadn't fully expressed symptoms at the beginning, and later in the experiment by surviving seedlings that grew a second flush that, with the naked eye, did not seem to be affected by the infection. This made these seedlings indistinguishable for MS from control seedlings. But in the middle, HS and MS classification accuracy was comparable.

II. Results

4. On average, classification accuracy using HS data is 10 percent points higher than the classification accuracy using MS data.
5. The difference in classification accuracy was mainly caused by the difference in precision and recall (probability of detection) of non-symptomatic seedlings (group 1). Where the precision using MS was 20 percent points lower than HS, and the probability of detection was 30 percent point lower. On a related note for early detection, pre-symptomatic detection rate for HS was 98% and 85% for MS.
6. Over time, the HS achieved a higher classification accuracy than MS, except for the last date.

DO's & DON'Ts:

Do use hyperspectral imaging to get on overall higher classification accuracy (1&4).
Do use hyperspectral when you need to identify individuals that are pre- or non- symptomatic (2&5). However, an early detection rate of 85% for MS still high and it might be an acceptable trade off for the cost price difference of ~$95,000.
Don't use hyperspectral imaging when individuals show symptoms (5). The precision and recall for groups 2 and 3 are comparable for MS and HS. Thus if the goal is to automatically detection symptomatic seedlings, use a multispectral sensor.

Acknowledgements