Virginia Information Technologies Agency

Statewide Land Cover 2015

Quality Assessment / Quality Control

Accuracy Report

Statewide

Prepared for:

Virginia Information Technologies Agency

December 9, 2016

Prepared by:
Sanborn Solutions
1935 Jamboree Drive
Suite 100
Colorado Springs, CO 80920

Table of Contents

List of Figures and Tables

VITA Land Cover QA\QC

1   Introduction

This final report for the VITA Land Cover QA/QC project covers the data for the state of Virginia. The total area of this delivery is 42,635.35 square miles.

In the six previous reports, Sanborn used a combination of image analysis and manual photo interpretation to select random samples, stratified throughout the extent of the individual six delivery areas, and labeled the samples according to the VITA classification scheme. Sanborn delineated samples with reference to the source imagery where needed in order to more accurately represent the landscape. Most samples used a buffered point with an area of 25 m^2. Those that were manually delineated used approximately the same area. Since an equal area was used for all points, each point could be taken as an equal representation of the classification.

The accuracy assessment points from the previous six delivery areas were used to sample the updated land cover data and generated statistics that quantify its accuracy.

The overall accuracy target for this project is 95%. Targets for individual classes are listed in Table 2. The Deterministic Assessment resulted in an overall accuracy statistic of 89%, while the Fuzzy Assessment resulted in an overall accuracy statistic of 93%.

A total of 9,418 points were chosen at random throughout the individual delivery areas, stratified by and proportional to the area of each class. A minimum point count of 50 was targeted per class per delivery area; when all points were tallied, Barren was the class with the smallest area, and a total of 278 points were assessed. The Forest class had the largest area, and 4,248 points were chosen. As in the previous reports, the wetland classification was collapsed to one classification which therefore includes Emergent Wetlands (92) and Woody Wetlands (91). The Impervious classification was collapsed into one classification for the accuracy assessment which therefore includes Impervious Extracted (21) and Impervious Datasets (22). The Forest and Tree classifications were treated as "fuzzy" classifications which mean a forest call also had a fuzzy tree call and a tree call also had a fuzzy forest call. The stratification and proportions are provided below in Table 1.

VITA Statewide Point Stratification

Class	Value	Sq. Miles	Proportion	Points
Open Water	11	1908.5	0.04476	573
Impervious	21	1562.6	0.3665	454
Barren	31	114.2	0.00268	278
Forest	41	23990.73	0.56270	4248
Tree	42	2560.19	0.06005	384
Shrub / Scrub	51	279.82	0.00656	403
Harvested / Disturbed	61	961.31	0.02255	431
Turfgrass	71	2279.96	0.05348	547
Pasture	81	4946.68	0.11602	968
Cropland	82	2151.14	0.05045	626
Wetlands	91	1880.17	0.04410	506
Total		**42635.35**	**1**	**9418**

Table 1: Point Stratification for VITA Statewide

Once the points were chosen for each individual delivery area and labeled by a photo interpreter, our internal QA/QC process ensured that each sample in the accuracy assessment database was labeled and delineated properly. This involved a QC analyst checking approximately 25% of the random points. The accuracy assessment database, once complete, was used to sample the land cover data and to generate the statistics to quantify its accuracy. Those statistics were compared to per-class target accuracies (below).
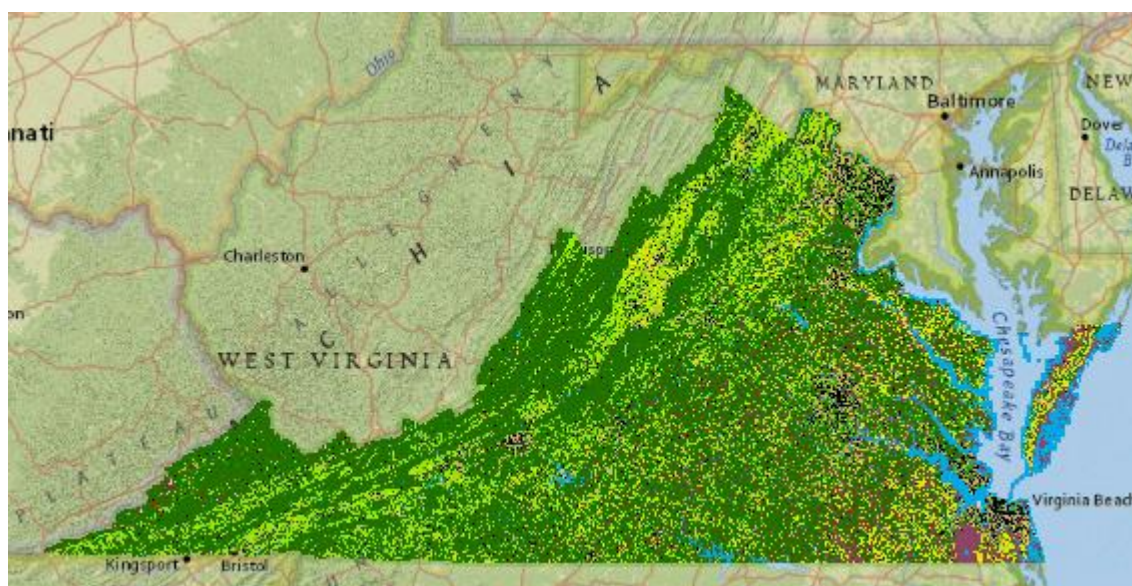

Figure 1: VITA Statewide Land Cover

	Land Cover		Minimum Mapping Unit	Resolution	Accuracy
Classification Categories	Herbaceous	Turf Grass	Less than 2 acres with land use exceptions	1 Meter	85%
	Impervious	Extracted Buildings, driveways, parking lots, roads ,etc	Match resolution	1 Meter	95%
		External Local & Statewide Impervious data	Road centerline dependent	1 Meter	95%
	Forest	Forest	1 acre w/ min width restrictions	1 Meter	95%
		Tree	Less than 1 acre	1 Meter	85%
		Harvested/Disturbed Forest	1 acre w/ min width restrictions	1 Meter	85%
	Scrub/Shrub	Scrub/Shrub	1 acre w/ min width restrictions	1 Meter	85%
	Agriculture	Cropland	1 acre w/ min width restrictions	1 Meter	85%
		Pastureland	1 acre w/ min width restrictions	1 Meter	85%
	Wetlands	NWI/Other	As defined by NWI and TMI	1 Meter	85%
	Barren	Barren	Higher than the resolution	1 Meter	85%
	Water	Water	Higher than the resolution	1 Meter	95%

Table 2: Classification scheme with class-specific parameters as defined in VGIN_LandCover_TechnicalPlanOfOperations_v7_ 20160506.pdf.

2 Accuracy Statistics

A deterministic matrix is presented first, using the complete VITA classification scheme. For the deterministic matrix, map classes were compared with only the primary photo interpretation. Fuzzy classes were not used.

2.1 Deterministic Accuracy Matrix

VITA Statewide Deterministic Accuracy Matrix

Accuracy Assessment Labels

Map Labels	Water	Impervious	Barren	Forest	Tree	Scrub/Shrub	Harvested/Disturbed Forest	TurfGrass	Pasture	Cropland	Wetlands	Total	Deterministic User's Accuracy
Water	556											**556**	100%
Impervious		438	6	4	6	2		9	4			**469**	93%
Barren	3	2	244	2		21	2	1	7			**282**	87%
Forest	2	2	1	4059	66	28	35	5	31	9	4	**4242**	96%
Tree	4	3	2	135	281	13	1	51	21	4		**515**	55%
Scrub/Shrub		1	3	14	8	248	18	1	25		2	**320**	78%
Harvested/Disturbed Forest		1	8	16	3	36	367		6	1	16	**454**	81%
TurfGrass	1	6	8	16	15	33	7	456	85	21		**648**	70%
Pasture			6	2	5	18	1	22	760	119	2	**935**	81%
Cropland		1				4		2	29	472	3	**511**	92%
Wetlands	7										479	**486**	99%
Total	573	454	278	4248	384	403	431	547	968	626	506	**9418**	
Deterministic Producer's Accuracy	97%	96%	88%	96%	73%	62%	85%	83%	79%	75%	95%		
Kappa	96.8%	96.3%	87.4%	91.9%	71.6%	60.2%	84.4%	82.1%	76.1%	74.0%	94.4%		

Deterministic Overall Accuracy	88.8%
Kappa	85.3%
N	9418

2.2 Fuzzy Accuracy Matrix

In most classification schemes, the current scheme included, there are instances in which a land cover feature may fall into more than one class, depending on the interpreter. This is known as Fuzzy Set Theory, or Fuzzy Logic, and is based on the idea that the landscape, or land cover, is not a set of discrete land cover types, but a continuum. It recognizes that, on the margins of classes that divide a continuum, an item may belong to both classes (Congalton and Green, 2009)[2]. For the VITA Land Cover project, examples of potentially fuzzy classes, or those with potential significant overlap in interpretations include: Scrub/Shrub and Forest, Scrub/Shrub and Harvested/Disturbed, Turfgrass and Pasture, Pasture and Cropland (especially when young or in a fallow or pre-planted state), and Forest and Woody Wetlands. Other specific instances might be ambiguous as well.

During the photo interpretation process, the photo interpreter made a Primary call, or the most likely label for the point. If there were other possible labels for the point, they were listed as Fuzzy1 and Fuzzy2. During compilation of the fuzzy matrix, if any of the photo interpreted labels matched the map label, the point was considered correct.

VITA Statewide Fuzzy Accuracy Matrix

Accuracy Assessment Labels

Map Labels	Water	Impervious	Barren	Forest	Tree	Scrub/Shrub	Harvested/Disturbed Forest	TurfGrass	Pasture	Cropland	Wetlands	Total	Fuzzy User's Accuracy
Water	556											556	100%
Impervious		440	5	4	5	2		9	4			469	94%
Barren	3	1	246	2		21	1	1	7			282	87%
Forest	2	2	1	4143		24	26	5	28	7	4	4242	98%
Tree	4	3	2		429	10	1	43	19	4		515	83%
Scrub/Shrub		1	3	11	5	277	11	1	9		2	320	87%
Harvested/Disturbed Forest		1	3	9	2	31	394		5	1	8	454	87%
TurfGrass	1	6	7	16	15	31	7	467	77	21		648	72%
Pasture			3	1	1	13	1	12	850	52	2	935	91%
Cropland		1				4		2	20	484		511	95%
Wetlands											486	486	100%
Total	566	455	270	4186	457	413	441	540	1019	569	502	9418	
Fuzzy Producer's Accuracy	98%	97%	91%	99%	94%	67%	89%	86%	83%	85%	97%		
Kappa	98.1%	96.5%	90.8%	98.1%	93.5%	65.9%	88.8%	85.5%	81.6%	84.2%	96.6%		

Fuzzy Overall Accuracy	**93.1%**
Kappa	**91.0%**
N	**9418**

3   Examples of Common Sources of Confusion

Referring to the fuzzy matrix, there were several common sources of confusion. The photo interpreted labels in the following examples were given by Sanborn; the Map labels were from the classification as provided by WorldView Solutions.

1. Scrub Shrub (Photo Interpreted) vs. Harvested / Disturbed (Map)
2. Scrub Shrub (Photo Interpreted) vs. Turfgrass (Map)
3. Turfgrass (Photo Interpreted) vs. Tree (Map)
4. Pasture (Photo Interpreted) vs. Turfgrass (Map)
5. Cropland (Photo Interpreted) vs. Pasture (Map)

Below are some visual examples of these confusions:
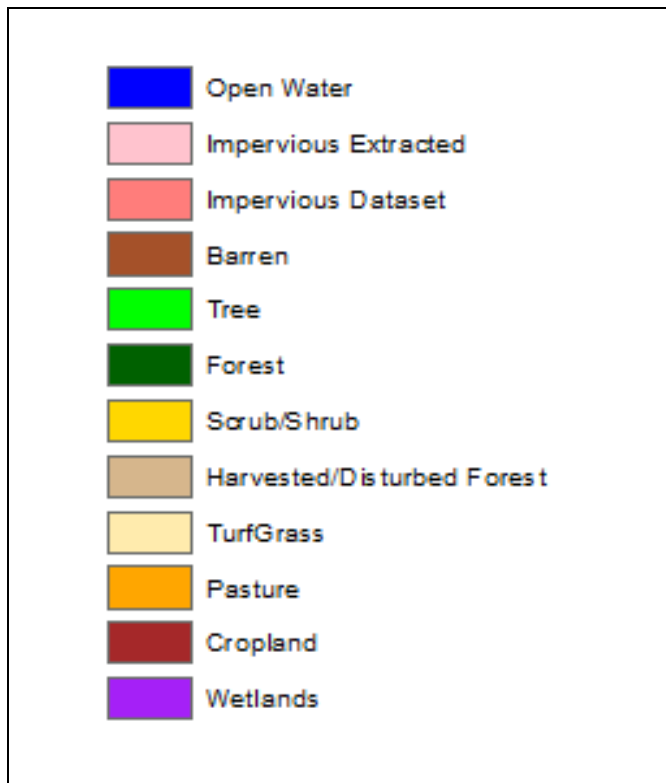
Table 3: Classification Legend

| 2011 VGIN Imagery | Classification – Beige is Harvested / Disturbed. |
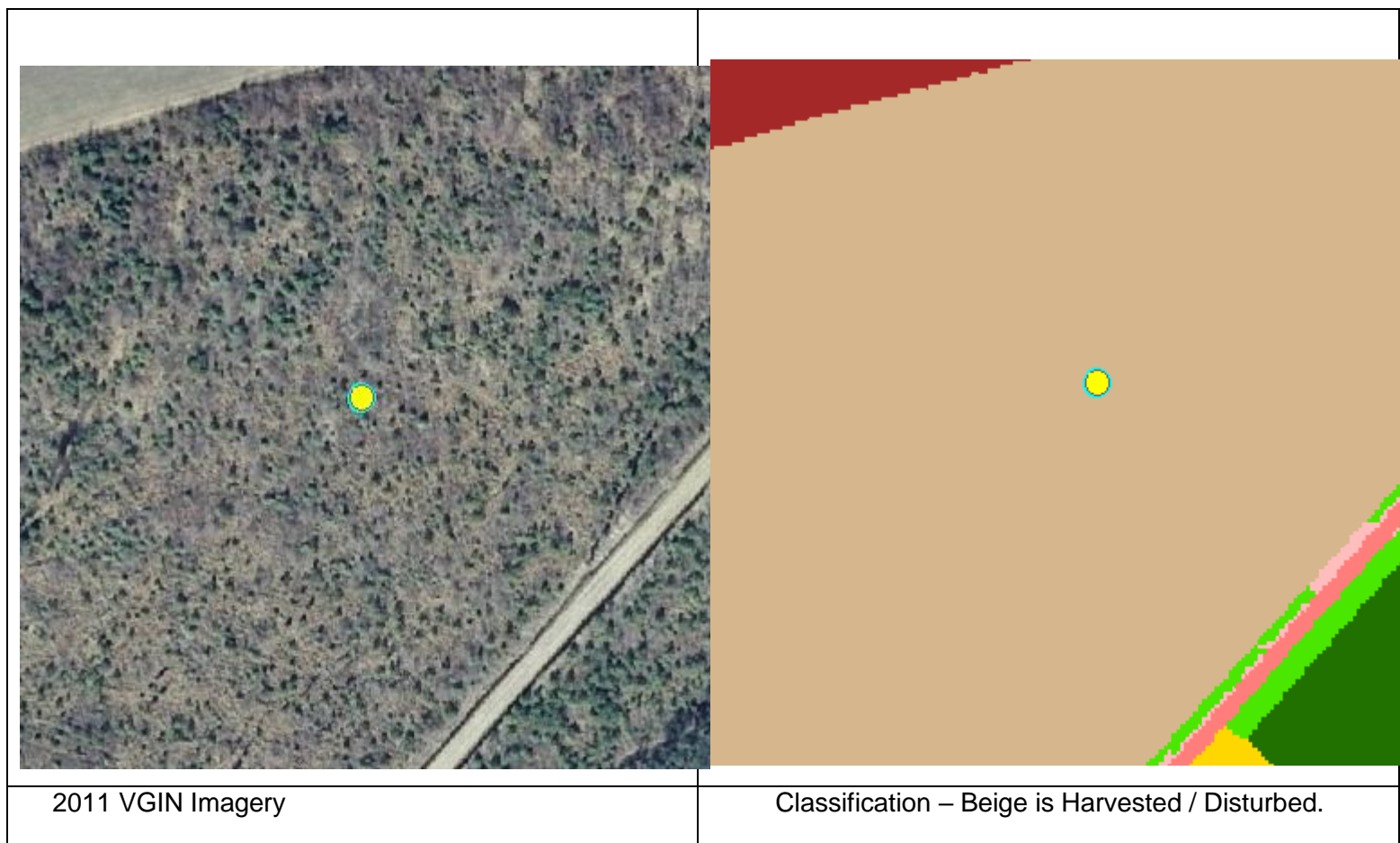
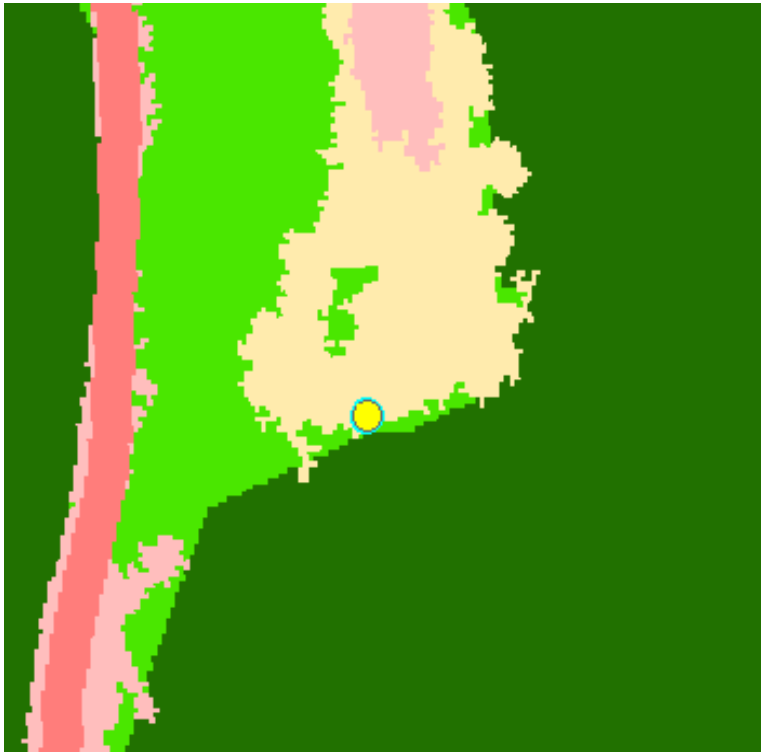**Figure 2:** Scrub Shrub vs. Harvested / Disturbed Confusion

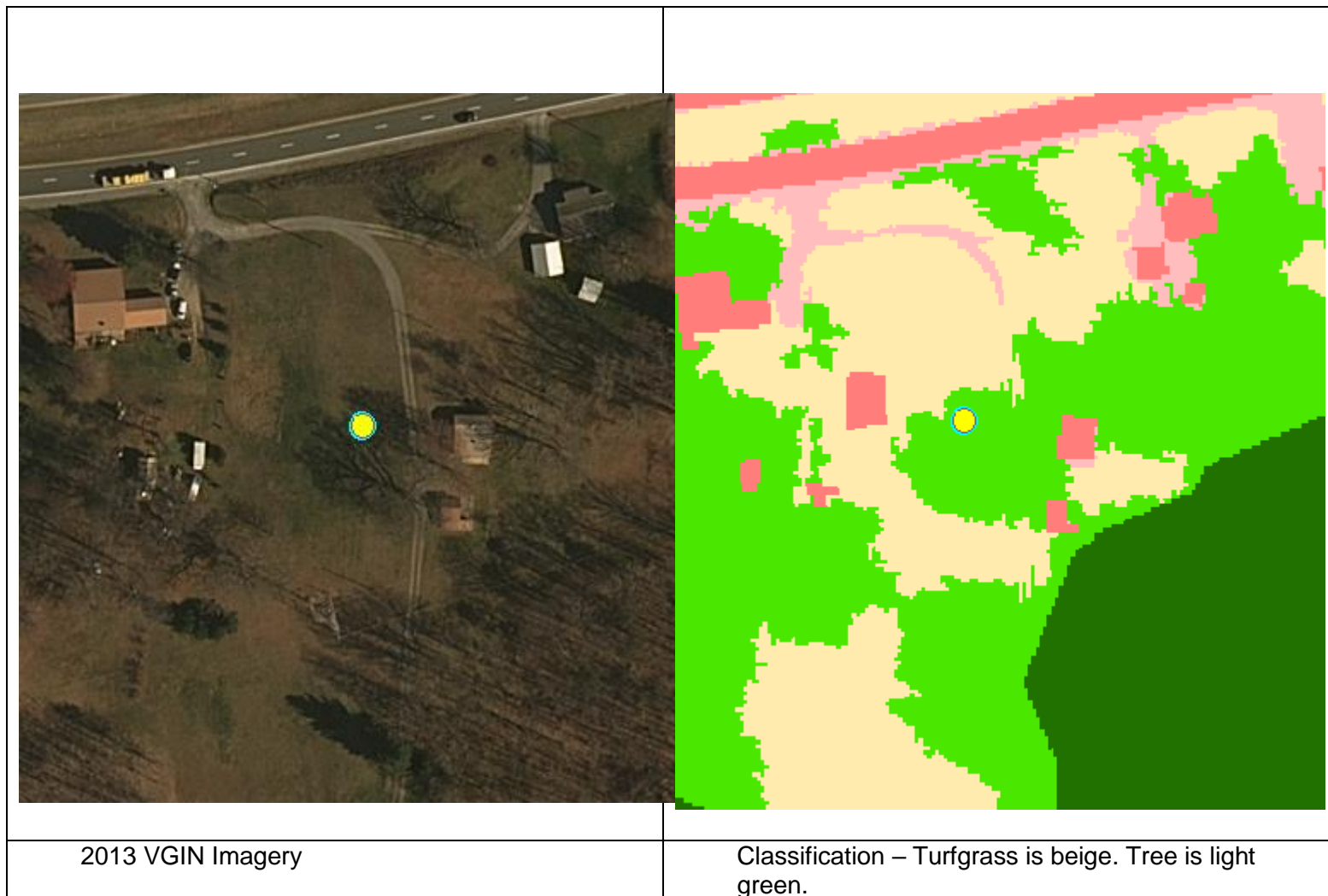|	
2015 VGIN Imagery	Classification – Tree is light green. Turfgrass is beige.

Figure 3: Scrub Shrub vs. Turfgrass Confusion

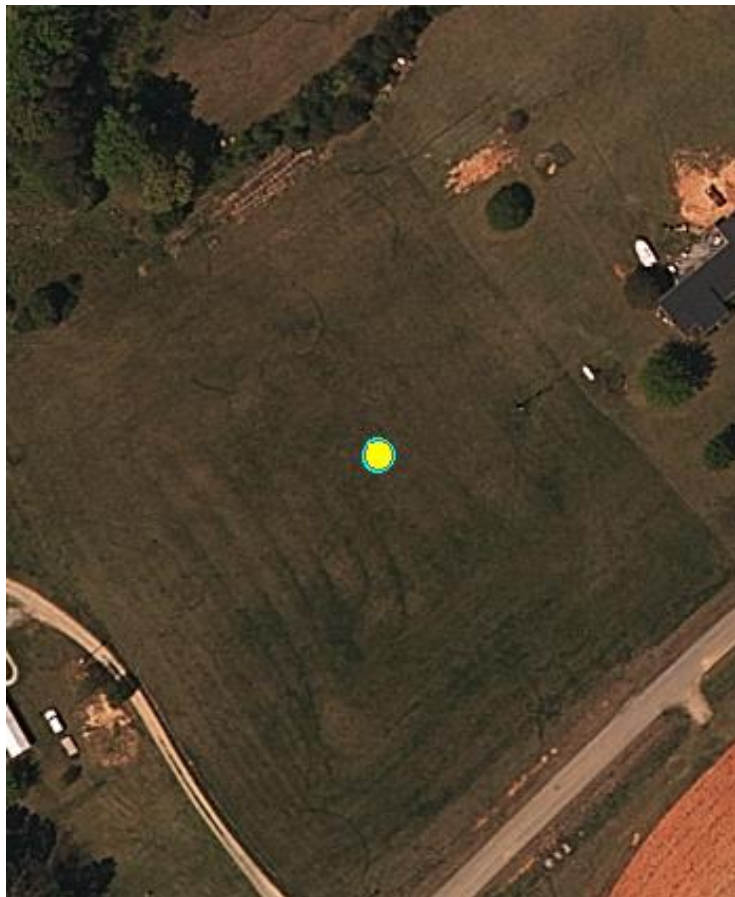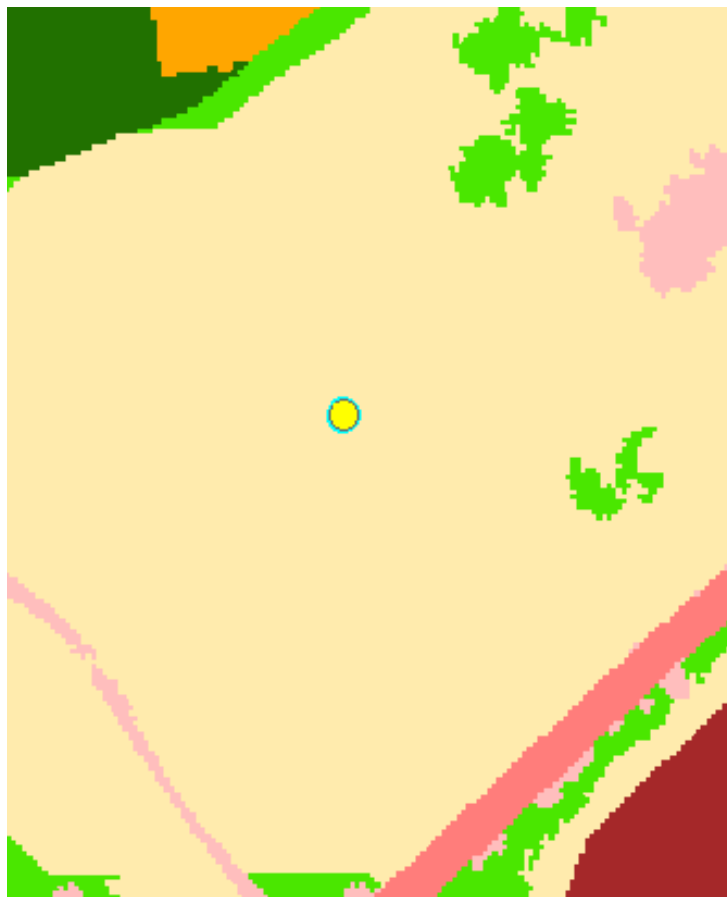| 2013 VGIN Imagery | Classification – Turfgrass is beige. Tree is light green. |

Figure 4: Turfgrass vs. Tree Confusion

| 2015 VGIN Imagery | Classification – Turfgrass is beige. |

Figure 5: Pasture vs. Turfgrass Confusion

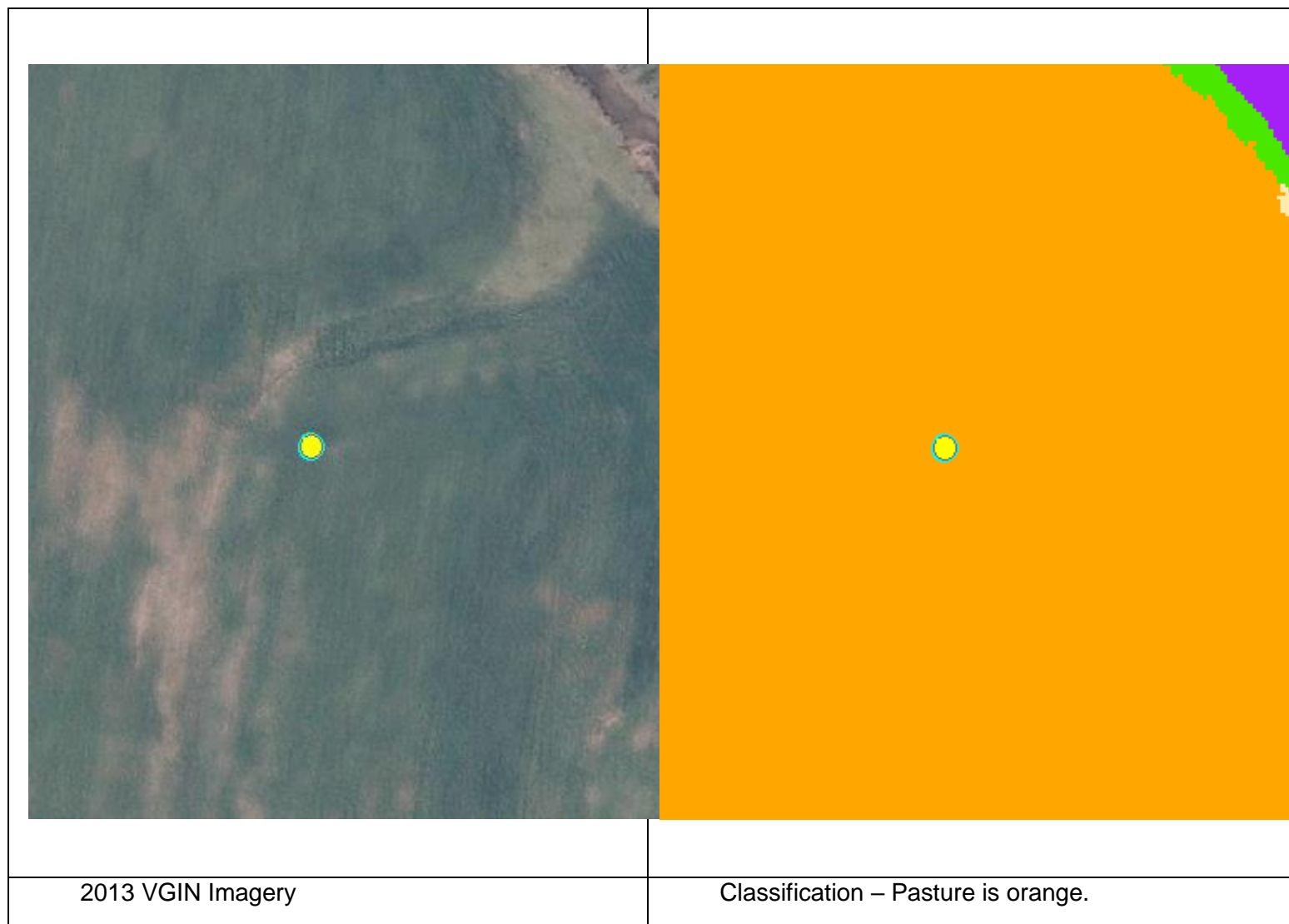| 2013 VGIN Imagery | Classification – Pasture is orange. |

Figure 6: Cropland vs. Pasture Confusion

4 References

Congalton, Russell G., Kass Green, Assessing the Accuracy of Remotely Sensed Data - Principles and Practices Second edition (2009) CRC Press, Taylor & Francis Group, Boca Raton, FL 978-1-4200-5512-2 183 pp.