“Ground-False”:

How Inaccuracy enters into Ground-Truthing

Robert Roth
reroth@wisc.edu

Introduction

- **Any remote sensing application is a very involved process, composed of many complicated and intricate steps**

  * e.g. data acquisition, data preparation, data processing, final output

- **One topic that is often over-looked, however, is the procedure for reference point collection**

  * Incredibly important because it is how we determine if our data processing technique actually worked!

Introduction

- **Not much attention is paid to the field collection of ground-truthing in mainstream RS literature or in the classroom**

  * Typically, there is only a discussion of how these reference points can be used (i.e. confusion matrices)

  * The validity of ground-truth is rarely questioned

Introduction

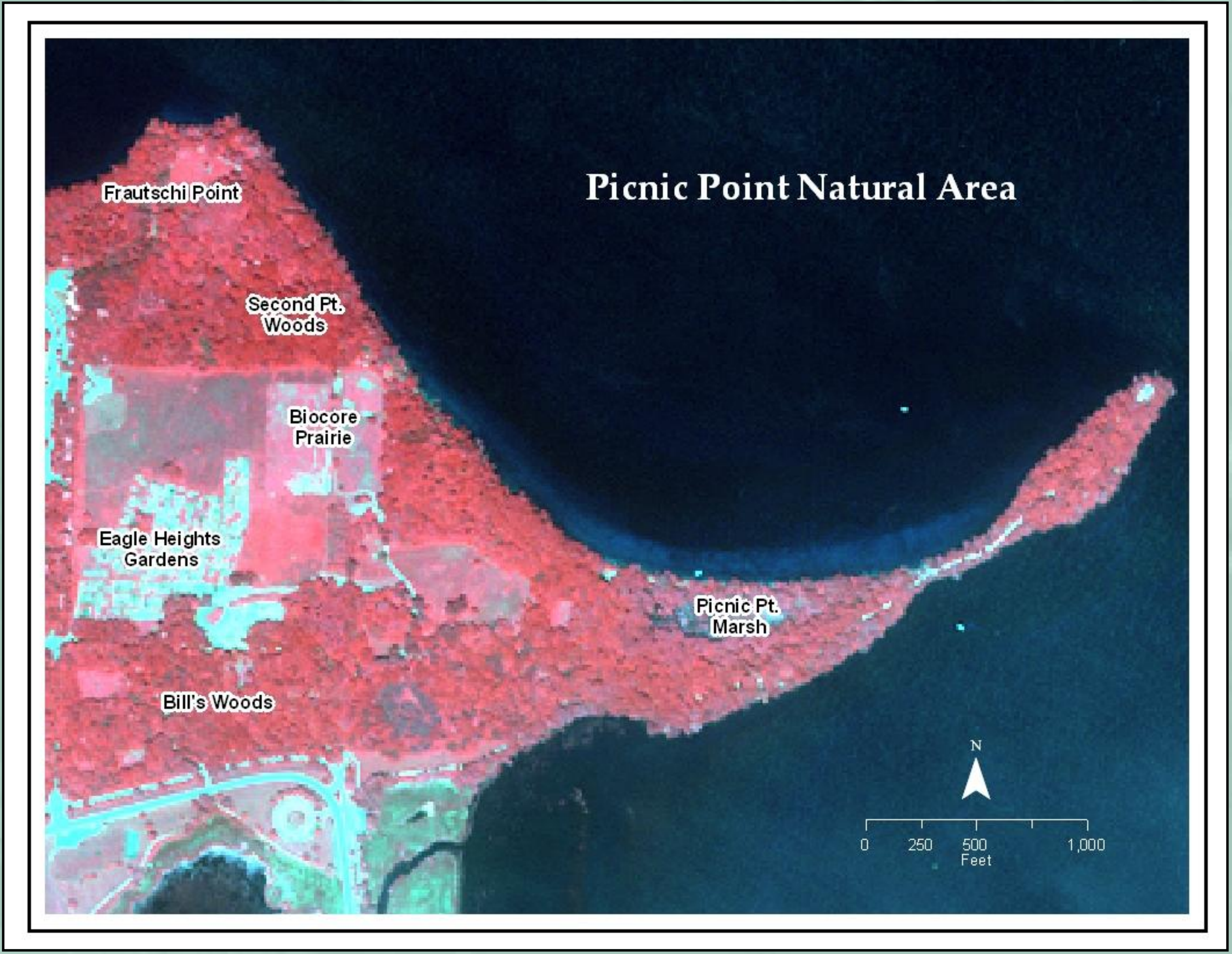
- Project goals:

1. Work through a case study of reference data collection

2. Use this experience to discuss ways in which inaccuracy can enter into ground-truth

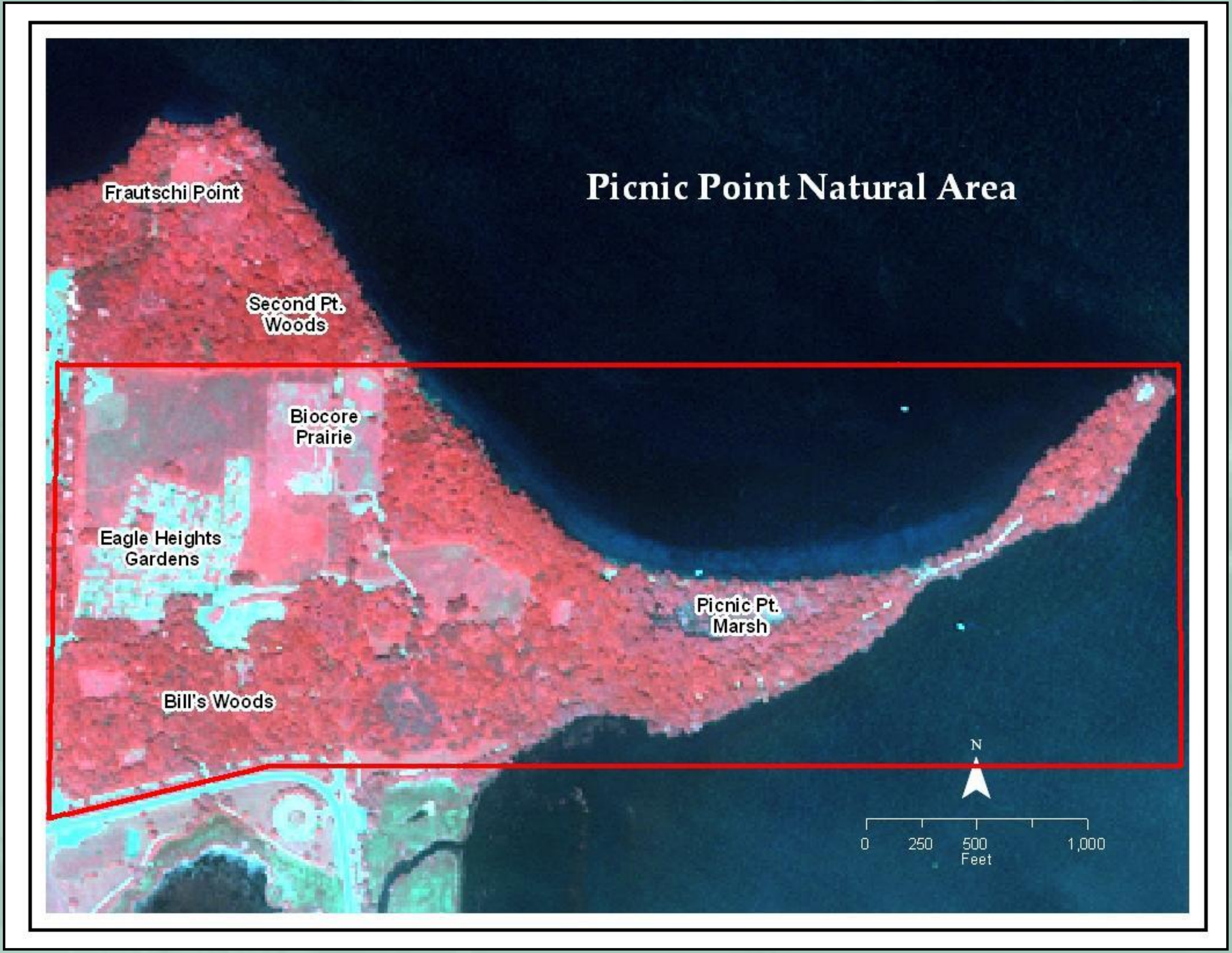
Methodology

Study Area: Picnic Point



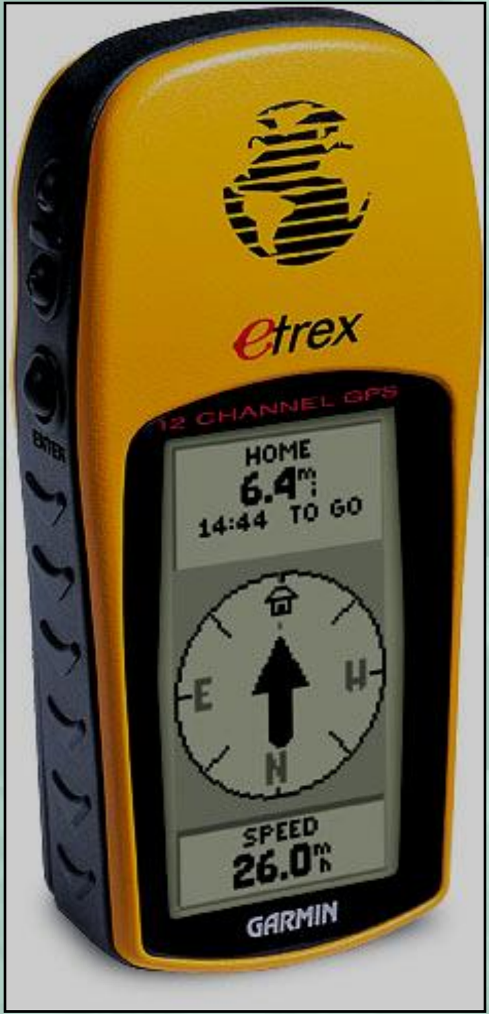
Methodology

Study Area: Picnic Point



Methodology

Data Collection Instrument:



- Garmin Etrex Vista GPS Unit

  * 15 foot accuracy with clear skies

  * Reads in LAT/LONG
    - degrees and decimal degrees

  * Built in compass & barometer

- GPS sampling is much different for RS applications than cartographic applications

Methodology

Classification System:

- The assumed end result for this ground-truth is the assessment of accuracy of a land cover classification

- Adopted a “classic” land cover classification scheme (from Lillesand et al.)

* Slightly altered based on study area

* Definitions based for view from the ground, not visual interpretation

Methodology

Classification System:

1. **Grass** – any short stemmed plant, not exclusive to the common lawn grass
2. **Garden** – exclusively the area within the Eagle Heights gardens
3. **Evergreen** – any isolated tree or continuous woods that produces evergreen canopy
4. **Deciduous** – any isolated tree or continuous woods of deciduous forest
5. **Marsh** – any grassland or forest with standing water
6. **Water** – any open water source
7. **Other** – any anthropogenic feature within the natural area

Methodology

Sampling Strategy:

- There are two general goals that must be attained when considering the appropriate sample strategy for a remote sensing application

1. A distribution of points that is representative of the entire study are
2. A large sample size within each determined classification category

Methodology

Sampling Strategy: Four Methods

1. **Random Sampling** - the selection of coordinates for sampling are generated from a random digits table across the entire study area

2. **Systematic Sampling -** a consistently spaced set of coordinates is produced for sampling by imposing a uniform grid upon the study area

3. **Stratified Sampling** - a minimum number of reference points are selected from each category

4. **Cluster Sampling -** clusters of sample points are gathered within several polygons of standard size that are placed randomly across the landscape.

Methodology

Sampling Strategy:

- In specific cases when no single strategy is perfect or two separate strategies are equally sufficient, a hybrid sampling scheme is often adopted

- In order to attain the original two goals, I used a *systematic-stratified hybrid* scheme

  * The initial systematic sampling will provide an even distribution of points

  * The subsequent stratified hybrid will ensure that each category has a minimum of 50 points

Methodology

Sampling Strategy: Systematic

- Points evenly spaced 0.0005 degrees apart longitudinally and latitudinally

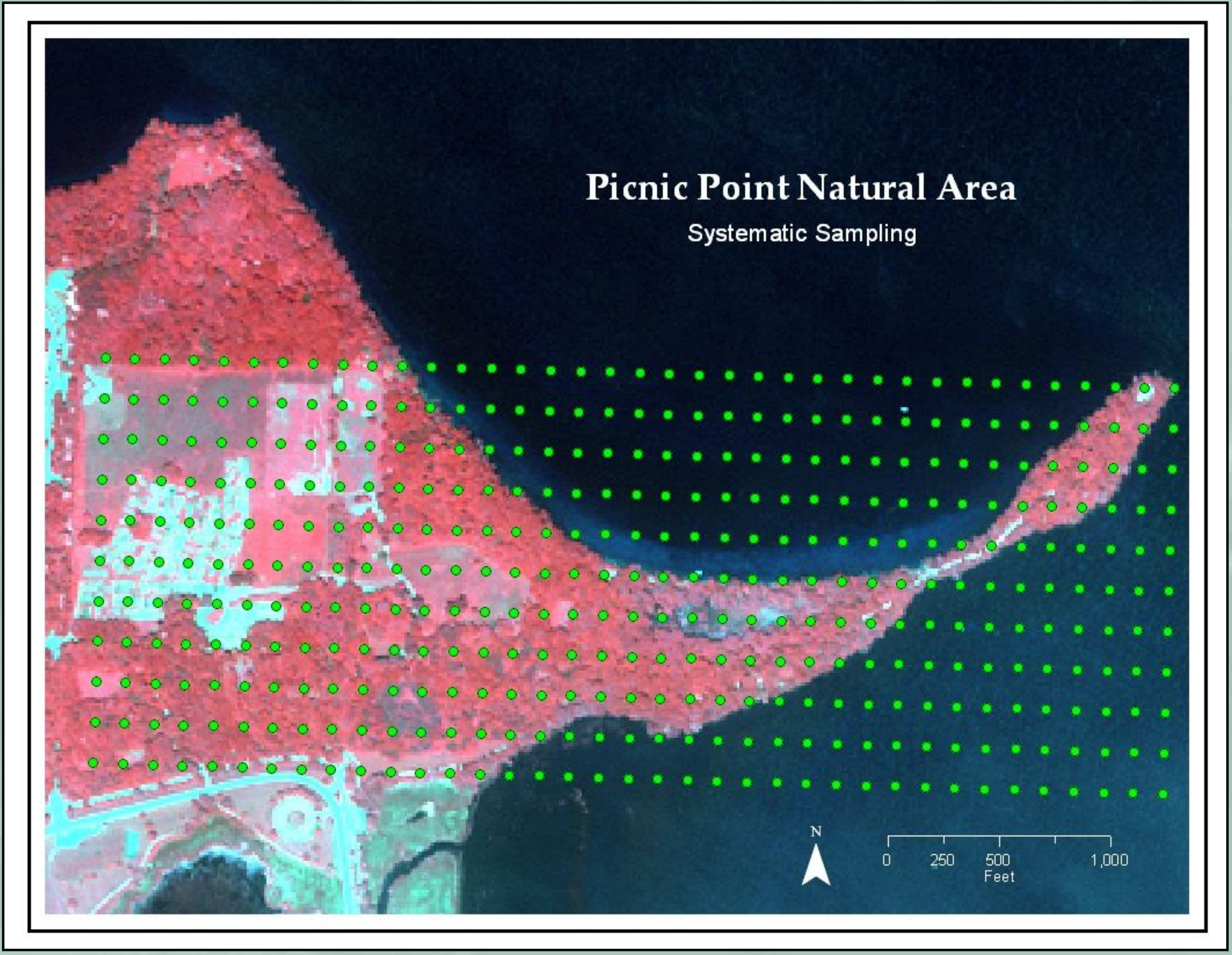
* The result is a grid of points that are spaced approximately 150’ along the horizontal and 200’ along the vertical

- The systematic sampling produces 407 points

* 11 rows and 37 columns

Methodology

Sampling Strategy: Systematic



Methodology

Sampling Strategy: Systematic

- The systematic aspect of the hybrid was very problematic
  * The transects cut through cover types that were difficult to traverse (thick undergrowth, marsh)
  * Issue of sampling Lake Mendota
  * Coordinates often in boundary areas that were difficult to classify
- Took twenty-five hours total, or three minutes per point

Methodology

Sampling Strategy: Stratified

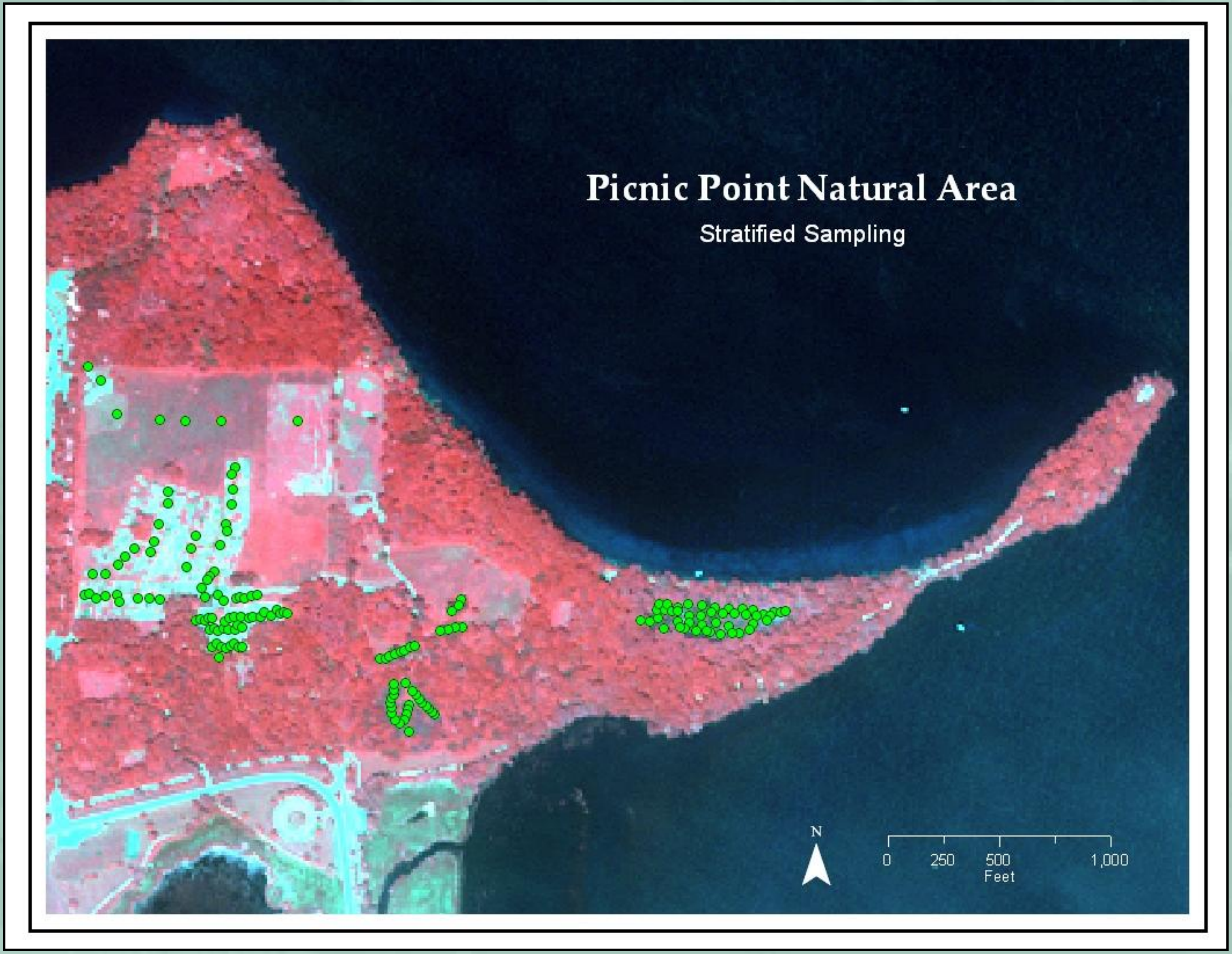
- The frequency per class of the initial sampling is as follows:

Water:	196
Deciduous:	122
Grass:	49
Garden:	17
Evergreen:	12
Marsh:	5
Other:	5

- To supplement the under-sampled categories, the stratified sampling approach is then implemented

Methodology

Sampling Strategy: Stratified



Methodology

Sampling Strategy: Stratified

- The stratified approach was much easier to implement
    * The points can be taken much closer together
    * The surveyor can collect points that unmistakably belong to a particular class

- It took three hours to collect 166 points, or about a point a minute
    * Makes sense why this is the method of choice for RS applications

Methodology

Sampling Strategy:

- The final frequency per class of the 573 points is as follows:

Water:	196
Deciduous:	122
Grass:	54
Evergreen:	51
Garden:	50
Marsh:	50
Other:	50

Methodology

Database Creation:

- **All 573 points must first be entered into an Excel workbook**
  * Then converted to a dBASE IV file that can be read by ArcCatalog

- **Then converted to a feature class in ArcCatalog**
  * The result is an XY shapefile with one useful column, the attribute or land cover type

Methodology

Registration:

- **The image used for classification is from IKONOS MSS taken in June 2002**

  * Multispectral resolution of four meters (which is roughly the accuracy level of the GPS unit)

- **The point shapefiles must be registered to the projection of the IKONOS MSS image**

  * UTM Zone 16

  * Resampling of units from degrees to meters

Methodology

Row/Col Conversion:

- The final step for reference point creation is to translate the locations of each XY point into terms of rows and columns of the IKONOS image

- Because I couldn't find a better solution, I had to manually generate the row/col values by comparing two windows
    * Incredibly time-consuming and error prone
    * There has to be a better way to do this!

Results

Unsupervised Classification:

Results

Contingency Table:

Reference Data

Classified Data

	Water	Deciduous	Evergreen	Grass	Garden	Marsh	Other	TOTAL
Water	194	2	0	0	0	0	0	196
Deciduous	0	94	8	1	0	2	0	105
Evergreen	0	8	35	1	0	3	0	47
Grass	0	13	7	47	11	10	4	92
Garden	0	0	0	2	35	1	5	43
Marsh	1	5	1	2	0	34	1	44
Other	1	0	0	1	4	0	40	46
TOTAL	196	122	51	54	50	50	50	573

Overall Accuracy
83.60%

Producer's Accuracy:

Water:	98.98%
Deciduous:	77.05%
Evergreen:	68.63%
Grass:	87.04%
Garden:	70.00%
Marsh:	68.00%
Other:	80.00%

User's Accuracy:

Water:	98.98%
Deciduous:	89.52%
Evergreen:	74.45%
Grass:	51.09%
Garden:	81.40%
Marsh:	77.27%
Other:	86.96%

Discussion

How Ground-Truth Becomes Ground-False:

1. The Collection Instrumentation

* “The Garmin Etrex Vista has an accuracy level of 15 feet” – lucky that this is only about a pixel for the image to be classified

* Severity depends on the scale

* Many features that are ground-truthed need secondary instruments (e.g. elevation)

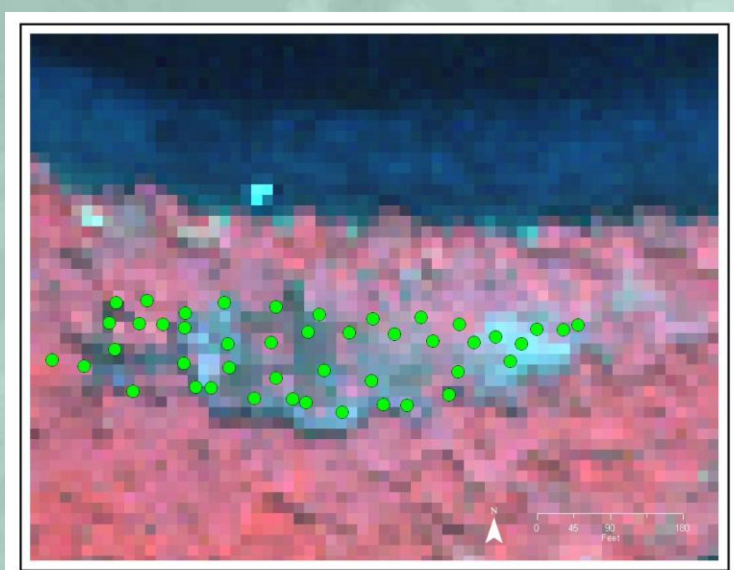
Discussion

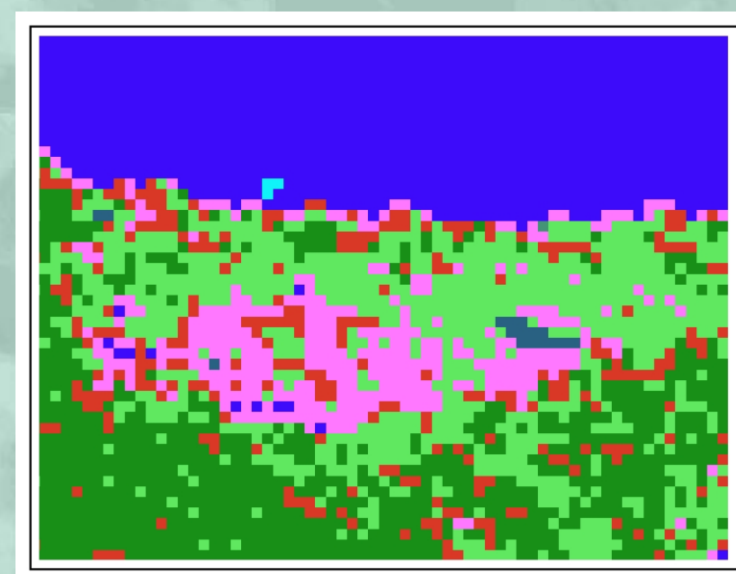
How Ground-Truth Becomes Ground-False:

2. The Classification Scheme

* A category is often defined by a visual interpreter

* In this case study, marsh was particular difficult to determine on the ground





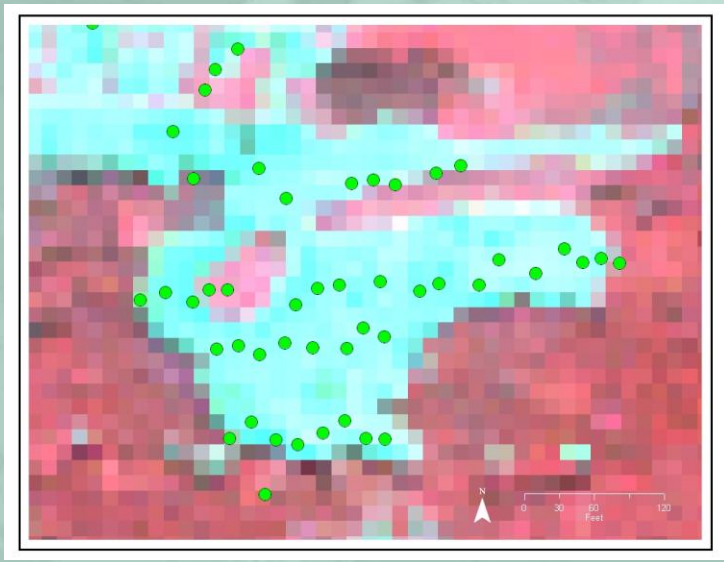
Discussion

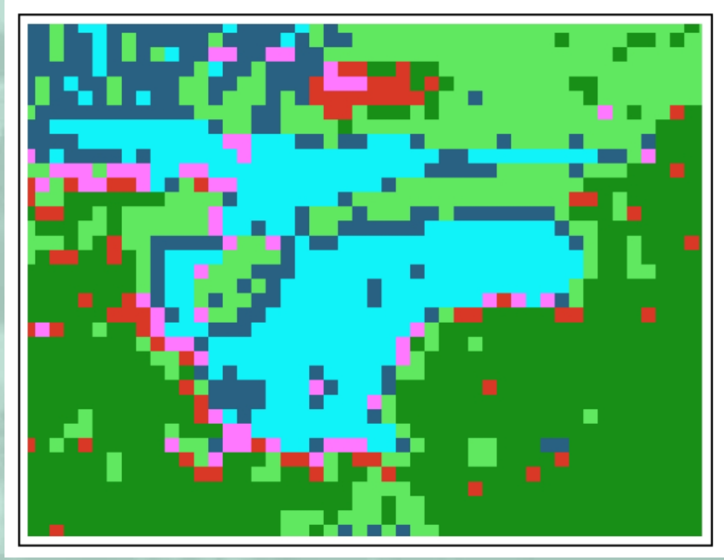
How Ground-Truth Becomes Ground-False:

3. Seasonality and Temporal Change

* An image is a static snapshot in time

* The area covered by the image is non-static, changing greatly from year-to-year, season-to-season and even from minute-to-minute





Discussion

How Ground-Truth Becomes Ground-False:

4. Database Entry

* Monotonous process of hand-entering the points into an Excel database

* Dealing with tens of thousandths of degrees, so even a slight mistake has a large effect on the spatial location

* Once examined in ArcMap, about 5% of the points had to be reentered because of spatial inaccuracy

Discussion

How Ground-Truth Becomes Ground-False:

4. Database Entry

* Monotonous process of hand-entering the points into an Excel database

* Dealing with tens of thousandths of degrees, so even a slight mistake has a large effect on the spatial location

* Once examined in ArcMap, about 5% of the points had to be reentered because of spatial inaccuracy

Discussion

How Ground-Truth Becomes Ground-False:

5. Registration

* The conversion from geographic space to UTM Zone 16 space is not exactly perfect

* Change of units from degrees to meters also introduces inaccuracy

Discussion

How Ground-Truth Becomes Ground-False:

6. Row/Col Conversion with Raster Images

* Matching the pixels between two windows proved to be one of the most difficult challenges of the entire project

* ArcMap and WiscImg stretched the image differently, so I had to match patterns, not specific colors

* I would guess the average error in corresponding locations would be at least one pixel, but likely multiple pixels for the whole reference data set

Conclusion

- Similar to inference statistics (Type II error), the sampled reference points only have a particular *explanatory power*

* The reference data is used to try to summarize the inaccuracy in the classification

* However, the inaccuracy within the reference data itself reduces the power or ability of the reference data to accuracy summarize this inaccuracy

* To improve the power, the sample size must be increased

Conclusion

- So should we just completely do away with ground-truthing all together?

* NO! It is still the best way to judge if our digital image processing technique was successful

* As good scientist, though, we need to be cognizant of the ways in which inaccuracy can be introduced into the reference data, and explore ways in which this inaccuracy can be reduced

Questions?

~ Thank you for your time,
Rob