

2019 STATEWIDE LAND USE MAPPING ACCURACY ASSESSMENT

After completing the final classification dataset, a comprehensive accuracy assessment was completed. On average, about one quarter of the independent ground truth samples were set aside for this purpose. A stratified random sampling method was used for accuracy assessment sample selection. The datasets were stratified by land cover type and county boundary. Prior to the accuracy assessment, the validation dataset was cleaned so that each data point corresponded to a single crop at a single point in time in a single field. In the WY 2019 analysis, 9,003 samples were used for accuracy assessment. These sites were not used to train the classification algorithm and therefore represent unbiased reference information.

Accuracy assessment can be divided into three components:

1. Reference data sampling design – how much reference data is collected, when and where
2. Reference data response design – how reference data is collected
3. Analysis – how the reference data is used to determine accuracy and how that accuracy is expressed

In this section, Land IQ's approach to all three of these components is described, and results of the accuracy assessment are provided.

SAMPLING DESIGN

In conventional accuracy assessment theory, the minimum number of samples corresponding to a specific accuracy level is calculated; however, this calculation does not take the size of the area being mapped into account. This method of minimum sample size determination is not always applicable to statistics with a spatial component. Therefore, rather than calculating the minimum number of samples needed for the reference data set, then dividing it up proportionately or equally among strata (crop types and cropped areas [e.g., counties]), Land IQ collects reference data by region or county and then partitions it into training and reference data sets for each crop based on an approximate 75%-25% split, respectively.

There are several reasons for this approach, but the main reason is that it is far more efficient to collect both training and validation data simultaneously when the reference data is collected by ground survey. This approach results in both types of data being concentrated where crops are concentrated (e.g., in counties with high acreage of agricultural crops). This approach also results in more data (for both training and testing) being collected for predominant, high acreage crops. The result of this approach repeated over multiple years is that representative reference data is accumulated, in some cases, beyond the minimum number of samples needed for some crops but can be lacking for other, less common crops. For this reason, a portion of data collection efforts are focused on crops for which data lacks. These crops are typically lower acreage, annual crops.

RESPONSE DESIGN

Reference data can be collected by different means from different sources. Currently, Land IQ collects reference data for model training and validation from cropped areas in California by conducting on-the-ground "ground truth" survey. For the purposes here, ground truth data and reference data are synonymous.

Because the ground truth surveys are real-time (as opposed to using previously acquired data such as archival imagery) and require presence of staff, logistical considerations must be made. First, on-the-ground reference data surveys must be made when the crop is growing. This requirement introduces an element of timing, which is especially important for short-season crops. Second, because the area mapped in California is so large, knowledge of where some cropped fields are, especially for minor crops, is approximate and often changes from year to year. In addition, because so many crops are

mapped that vary in acreage, some crops for which there is little existing reference data are prioritized for reference point data collection.

For these reasons, Land IQ uses the basic concepts of sampling design to achieve independent and random samples in addition to considering criteria to prioritize reference point data collection:

- Confidence level – Crops with estimated lower accuracy and confidence levels from the previous year’s mapping effort are prioritized for ground truth data collection.
- Peak date – Remote sensing analysis is used to find dates of peak reflectance in fields to determine the seasonality of crops and help optimize timing of ground truth data collection around peak growing seasons.

ANALYSIS

Uncertainty in crop classification is related to two issues: accuracy and precision.

ACCURACY

Accuracy is a relative measure of the exactness of an estimate and accounts for systematic errors. Therefore, an accurate estimate does not systematically over- or underestimate the true value. Map accuracy can be quantified by creating an error matrix (also commonly called a confusion matrix), which compares the map classification with a reference classification.

The underlying principle of the accuracy assessment is that it compares the mapped land classification to reliable reference data, collected through sample-based approaches, as described above. The objective of a validation data set, therefore, is to provide a statistically sound estimate of the accuracy of the output map based on an independent reference information source. The accuracy of a map is assessed by measuring the degree of agreement between the output map and the validation data set. An error matrix can be generated that compares the pixels or polygons in the resulting classification map to the known reference points. From this matrix, overall accuracy and accuracy of each class can be determined.

There are three measures of accuracy that can be determined from an error matrix:

1. Overall accuracy
2. Producer’s accuracy (omission error)
3. User’s accuracy (commission error)

Typically, accuracy of remotely sensed maps is demonstrated using an error (or confusion) matrix (Table 5). Accuracy measures that can be derived from an error matrix are described below.

OVERALL ACCURACY

Overall accuracy is calculated by dividing the total number of correct pixels by the total number of pixels in the error matrix. In other words, the total number of correctly classified samples are divided by the total number of samples. It measures the accuracy of the whole map but does not refer to any individual classes. It is the probability that a randomly selected location on the map is correctly classified. Overall accuracy is sensitive to sample size and is thus more reliable in classes with larger samples. It is the sum of the major diagonal in an error matrix that runs from the upper left corner to the bottom right corner of the matrix.

PRODUCER’S ACCURACY (OMISSION ERROR)

Producer’s accuracy is also called the omission error and is described by the probability that a reference point is correctly classified. It indicates how well the area represented by the map can be classified. It is shown on the right side of the matrix.

USER'S ACCURACY (COMMISSION ERROR)

User's accuracy is the number of correct fields in an individual class divided by the number of pixels that were actually classified in that class. It is called user's accuracy because it is the probability that a field classified on the map actually represents that class on the ground. This measure is also called commission error and is reported at the bottom of the matrix.

PRECISION

Precision is related to the random error, which can be quantified by a confidence interval. A confidence interval gives a range that encloses the true value of an unknown fixed quantity with a specified probability. A precise estimate would thus have a small confidence interval.

RESULTS

Accuracy was assessed based on both the DWR standard legend and Land IQ Legend. As the level of detail of these legends differed somewhat, the accuracy, or ability to correctly determine classifications, differs as well.

OVERALL ACCURACY

The overall accuracy for WY 2019 crop mapping statewide was 98.1% at the DWR legend level and 96.9% at the Land IQ legend level (Table 1).

Table 1. WY 2019 Overall Accuracy of Statewide Land Use Mapping

Crop Legend	Overall Accuracy (%)
DWR	98.1
Land IQ	96.9

The error matrices for crops classed by the DWR legend and the Land IQ legend (Tables 5 and 6 at the end of this report) show overall accuracy as well as omission and commission error, by crop class (in acres).

ACCURACY BY CROP CLASS

Accuracy was calculated for each crop class (number of correct acres divided by total acres in each crop category) for both DWR and Land IQ legends (Tables 2 and 3). Some land cover types (e.g., apples, avocados, bush berries, carrots, cole crops, dates, kiwis) are not included in the accuracy assessment due to insufficient data. In these cases, there were either no or less than five samples available for accuracy assessment. It is notable that some crops are missed; it is challenging to detect every crop instance because satellite data are intermittent and cropping is rotational and, in some cases, short term. For this reason, available data will not always align well with rotational crop timing. However, most missed crops are short season in nature and therefore have a smaller impact on total water use analysis. In total, the multi crop resolution of mapping data in WY 2019 captures a vast majority of the cropping year-round in the state, allowing data users to characterize crop production and water use more accurately. It should be noted that young perennials, while a smaller class, are challenging to detect with remote sensing approaches and can be confused with fallow until features are detectable. This is particularly true in years when higher resolution (e.g., NAIP) image resources are not available. NAIP imagery was not available in WY 2019, but it is expected that many young perennial features undetectable in 2019 will be detected in 2020 with higher resolution data availability.

Table 2. WY 2019 Statewide Land Use Mapping Accuracy by DWR Crop Legend

Crop	Correct Acres	Total Acres	Accuracy
Citrus and Subtropical	5,972	5,991	100%
Deciduous Fruits and Nuts	86,814	86,868	100%
Field Crops	68,581	70,396	97%
Grain and Hay	29,537	31,487	94%
Pasture	41,443	42,309	98%
Rice	30,344	30,344	100%
Truck, Nursery and Berry Crops	53,519	54,468	98%
Unclassified	16,873	17,064	99%
Vineyard	16,983	17,008	100%
Young Perennials	64	858	7%

Table 3. WY 2019 Statewide Land Use Mapping Accuracy by Land IQ Crop Legend

Crop	Correct Acres	Total Acres	Accuracy
Alfalfa and Alfalfa Mixtures	28,763	29,295	98%
Almonds	49,731	49,768	100%
Apples	203	215	94%
Avocados	1,725	1,734	99%
Beans (Dry)	2,083	2,542	82%
Bush Berries	366	366	100%
Carrots	1,987	2,244	89%
Cherries	1,407	1,407	100%
Citrus	2,336	2,387	98%
Cole Crops	8,546	9,309	92%
Corn, Sorghum and Sudan	37,746	38,537	98%
Cotton	16,929	17,090	99%
Dates	647	647	100%
Flowers, Nursery & Christmas Tree Farms	722	741	97%
Grapes	16,983	17,008	100%
Kiwis	79	79	100%
Lettuce/Leafy Greens	9,142	9,972	92%
Melons, Squash and Cucumbers	2,958	3,373	88%
Miscellaneous Deciduous	214	214	100%
Miscellaneous Field Crops	4,081	4,355	94%
Miscellaneous Grain and Hay	29,537	31,487	94%
Miscellaneous Grasses	4,870	5,997	81%
Miscellaneous Subtropical Fruits	9	9	100%
Miscellaneous Truck Crops	4,879	5,006	97%
Mixed Pasture	6,193	7,017	88%
Olives	1,135	1,135	100%
Onions and Garlic	5,102	5,243	97%
Peaches/Nectarines	1,716	1,780	96%

Crop Name	Correct Acres	Total Acres	Accuracy
Pears	703	703	100%
Peppers	1,114	1,203	93%
Pistachios	14,114	14,114	100%
Plums, Prunes and Apricots	2,779	2,862	97%
Pomegranates	412	412	100%
Potatoes or Sweet Potatoes	3,035	3,192	95%
Rice	30,344	30,344	100%
Safflower	3,452	3,523	98%
Strawberries	1,438	1,466	98%
Sunflowers	3,985	4,349	92%
Tomatoes	12,069	12,354	98%
Unclassified Fallow	16,873	17,064	99%
Walnuts	15,359	15,393	100%
Young Perennials	64	858	7%

PRECISION BY CROP

Two-tailed confidence intervals (95%) were calculated using the method in Olofsson et al. (2014) for the commission error of each crop class and are shown in Table 4 in order of highest to lowest precision. As noted above, precision is related to the random error, which can be quantified by a confidence interval. A confidence interval gives a range that encompasses the true value of an unknown fixed quantity with a specified probability. A precise estimate would thus have a small confidence interval. For example, avocados were mapped at 98% accuracy with a confidence interval of plus or minus 1%. This means that 98% of the time, we are confident that the pear classification was between 97 and 99% correct.

As Table 4 shows, 17 crops were mapped with 100% accuracy and 0% confidence interval (100% confidence or precision). An additional seven crops were mapped at accuracies ranging from 97 to 99% with 100% confidence. Table 9 also shows that the number of ground truth points directly influences the level of precision. As the number of ground truth points increases, precision (confidence) generally also increases and the confidence interval becomes smaller. Some crops are mapped with high accuracy with few ground truth points (such as kiwi fruit) because they are very distinct and relatively easy to distinguish from other crops. Other crops have a lower accuracy but relatively high precision (miscellaneous truck crops) because the number of ground truth points was relatively high. Some crops, such as apples, were mapped with high accuracy but lower precision (e.g., apples, dates) because of very few ground truth points.

Table 4. WY 2019 Statewide Land Use Mapping Accuracy and Precision by Crop

Crop Class	User's Accuracy (Number of correctly classified acres/total acres)	Number of Groundtruth (Reference Sample) Acres	95% Two- tailed Confidence Interval
Almonds	100%	49,768	0%
Bush Berries	100%	366	0%
Citrus	100%	2,387	0%
Corn, Sorghum and Sudan	100%	38,537	0%
Cotton	100%	17,090	0%
Dates	100%	647	0%
Grapes	100%	17,008	0%
Kiwis	100%	79	0%
Miscellaneous Subtropical Fruits	100%	9	0%
Olives	100%	1,135	0%
Pears	100%	703	0%
Pistachios	100%	14,114	0%
Pomegranates	100%	412	0%
Rice	100%	30,344	0%
Strawberries	100%	1,466	0%
Walnuts	100%	15,392	0%
Young Perennials	100%	857	0%
Cherries	99%	1,407	1%
Avocados	98%	1,734	1%
Miscellaneous Truck Crops	98%	5,006	0%
Safflower	98%	3,523	0%
Carrots	97%	2,245	1%
Onions and Garlic	97%	5,243	0%
Plums, Prunes and Apricots	97%	2,861	1%
Beans (Dry)	96%	2,542	1%
Miscellaneous Grain and Hay	96%	31,487	0%
Peaches/Nectarines	96%	1,780	1%
Alfalfa and Alfalfa Mixtures	95%	29,295	0%
Miscellaneous Field Crops	95%	4,355	1%
Mixed Pasture	95%	7,017	1%
Potatoes or Sweet Potatoes	95%	3,192	1%
Sunflowers	95%	4,350	1%
Cole Crops	94%	9,309	0%
Lettuce/Leafy Greens	92%	9,971	1%
Flowers, Nursery and Christmas Tree Farms	91%	741	2%
Tomatoes	91%	12,354	1%
Apples	90%	215	4%
Unclassified Fallow	90%	17,064	0%
Miscellaneous Grasses	86%	5,997	1%
Peppers	85%	1,202	2%
Miscellaneous Deciduous	83%	214	5%
Melons, Squash and Cucumbers	79%	3,374	1%

Table 5. Statewide Land Use Mapping Error Matrix by DWR Crop Legend (acres)

		Predicted										Total Acres	Correct Acres	Incorrect Acres	% Correct	Omission Error
		Citrus and Subtropical	Deciduous Fruits and Nuts	Field Crops	Grain and Hay Crops	Pasture	Rice	Truck, Nursery, and Berry Crops	Unclassified Fallow	Vineyard	Young Perennial					
Reference	Citrus and Subtropical	5,972	8	-	-	-	-	-	11	-	-	5,991	5,972	19	100%	0.3%
	Deciduous Fruits and Nuts	-	86,814	-	-	-	-	-	-	54	-	86,868	86,814	54	100%	0.1%
	Field Crops	-	8	68,581	111	408	-	1,283	6	-	-	70,396	68,581	1,815	97%	2.6%
	Grain and Hay Crops	-	-	57	29,537	323	-	655	914	-	-	31,487	29,537	1,949	94%	6.2%
	Pasture	-	-	-	727	41,443	-	-	139	-	-	42,309	41,443	866	98%	2.0%
	Rice	-	-	-	-	-	30,344	-	-	-	-	30,344	30,344	-	100%	0.0%
	Truck, Nursery, and Berry Crops	10	-	363	217	277	-	53,519	83	-	-	54,468	53,519	950	98%	1.7%
	Unclassified Fallow	-	-	-	112	79	-	-	16,873	-	-	17,064	16,873	192	99%	1.1%
	Vineyard	1	21	-	-	-	-	-	3	16,983	-	17,008	16,983	25	100%	0.1%
	Young Perennial	-	57	-	-	-	-	-	736	-	64	858	64	793	7%	92.5%
		Total Acres	5,983	86,908	69,001	30,705	42,530	30,344	55,457	18,765	17,038	64	356,794			
	Correct Acres	5,972	86,814	68,581	29,537	41,443	30,344	53,519	16,873	16,983	64	350,129				
	Incorrect Acres	11	94	420	1,167	1,087	-	1,938	1,892	54	-	6,665				
	% Correct	100%	100%	99%	96%	97%	100%	97%	90%	100%	100%	98.1%				
	Commission Error	0.2%	0.1%	0.6%	4.0%	2.6%	0.0%	3.6%	11.2%	0.3%	0.0%					

