



SMEX03 Land Use Classification Data: Oklahoma, Version 1

USER GUIDE

How to Cite These Data

As a condition of using these data, you must include a citation:

Jackson, T. and M. Cosh. 2009. *SMEX03 Land Use Classification Data: Oklahoma, Version 1*. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/ZBH68D2ZJ3JK>. [Date Accessed].

FOR QUESTIONS ABOUT THESE DATA, CONTACT NSIDC@NSIDC.ORG

FOR CURRENT INFORMATION, VISIT <https://nsidc.org/data/NSIDC-0365>



National Snow and Ice Data Center

TABLE OF CONTENTS

1	DETAILED DATA DESCRIPTION.....	2
1.1	Format	2
1.2	File and Directory Structure.....	2
1.3	File Naming Convention	2
1.4	Spatial Coverage.....	3
1.4.1	Spatial Resolution.....	3
1.4.2	Projection and Grid Description	3
1.5	Temporal Coverage.....	3
1.6	Parameter or Variable	3
1.6.1	Parameter Description	3
1.6.2	Sample Data Record.....	4
2	SOFTWARE AND TOOLS	4
2.1	Quality Assessment.....	5
3	DATA ACQUISITION AND PROCESSING.....	5
3.1	Data Acquisition Methods.....	5
3.2	Derivation Techniques and Algorithms.....	6
3.2.1	Processing Steps.....	6
4	REFERENCES AND RELATED PUBLICATIONS	6
4.1	Related Websites	6
5	CONTACTS AND ACKNOWLEDGMENTS	7
6	DOCUMENT INFORMATION.....	7
6.1	Publication Date	7
6.2	Date Last Updated.....	7

1 DETAILED DATA DESCRIPTION

1.1 Format

Land use classification data derived from Landsat 5 TM are provided in a flat, Band-Interleaved-by-Line binary file called SMEX03_OK_Land_Cover.bil. The Oklahoma study area image dimensions are 12,100 rows by 8201 columns. An Environment for Visualizing Images (ENVI) header file with geolocation information, called SMEX03_OK_Land_Cover.hdr, is also provided.

Additionally, in situ land cover classification data are provided in the Land_Cover.xls file. The file lists the initial surface conditions for each sampling site in the Oklahoma South and Little Washita regional study areas. Sampling sites are labeled by either OS (Oklahoma South) or LW (Little Washita) followed by the site number. For example, OS03 indicates the third sampling site in Oklahoma South study area. OS sampling site values range from 01-52. LW sampling site values include 02, 03, 04, 11, 12, 13, 20, 21, 22, 27, 28, 29, 31, 32, and 33. Also indicated in the Land_Cover.xls file is any noted change in surface condition, such as tillage, planting, or harvesting during the 2 - 18 July 2003 timeframe.

1.2 File and Directory Structure

All files are available on the HTTPS site in the https://daacdata.apps.nsidc.org/pub/DATASETS/AVDM/data/soil_moisture/SMEX03/Oklahoma/ancillary_data/landuse_classification/ directory:

1.3 File Naming Convention

Landsat 5 TM data are provided in a single binary file with a corresponding header file. These files are named according to the following convention and are further described in Table 1:

SMEX03_OK_Land_Cover.bil

SMEX03_OK_Land_Cover.hdr

Land_Cover.xls

Where:

Table 1 Description of File Name Variables

Variable	Description
SMEX03	Soil Moisture Experiment 2003
OK	Oklahoma
Land_cover	Parameter for this data set
.bil	Indicates that this is a Band-Interleaved-by-Line binary data file.
.hdr	Indicates that this is a header file.
.xls	Indicates that this is an Excel file.

1.4 Spatial Coverage

Southernmost Latitude: 33.6° N

Northernmost Latitude: 37.0° N

Westernmost Longitude: 99.3° W

Easternmost Longitude: 96.6° W

1.4.1 Spatial Resolution

The pixel size is 30 m by 30 m.

1.4.2 Projection and Grid Description

Universal Transverse Mercator (UTM), Zone 14, World Geodetic System 1984 (WGS 84)

1.5 Temporal Coverage

TM imagery was collected for 10 July 2003. In situ data were collected from 2 through 18 July 2003.

1.6 Parameter or Variable

The measured parameter for this data set is land use (vegetation) classification. Land use classification distinguishes between crop types, water, roads, and urban areas.

1.6.1 Parameter Description

The parameter range for various Oklahoma land cover types is described in Table 2. Valid parameter range values range from 0 to 8.

Table 2 Parameter Range and Description

Parameter Value	Description
0	Unclassified
1	Forest
2	Pasture
3	Winter Wheat/Bare
4	Alfalfa
5	Corn
6	Urban/Quarry
7	Soybean
8	Water

1.6.2 Sample Data Record

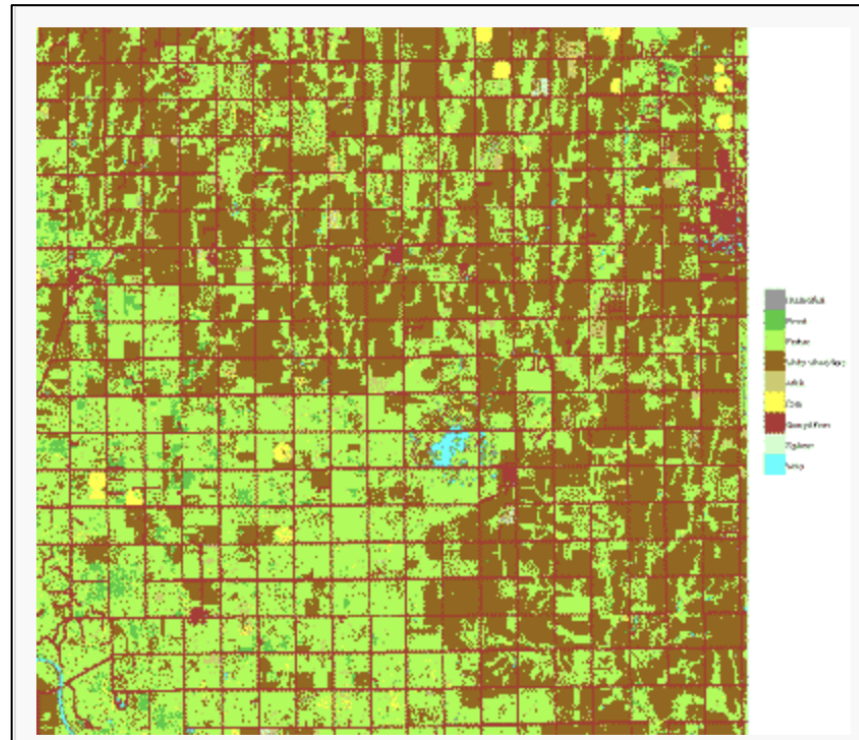


Figure 1 Land Use Classification Sample Data Record

2 SOFTWARE AND TOOLS

Tools appropriate for viewing these data include: [ArcView](#), [ENVI](#), or other similar visualization software packages; any text editor or Web browser; and any spreadsheet program.

2.1 Quality Assessment

Table 3 compares the accuracy of the in situ data with the accuracy of the image data obtained via satellite. The far right column lists the accuracy percentages of the in situ data when compared with the satellite image data. For example, of 581 classified forest pixels, only 566 of them are registered in the in situ data as forest, for a 97.4 percent accuracy. Accuracy percentages listed in the bottom row are the satellite image data compared to the in situ data. For example, of the 818 pixels that were deemed to be winter wheat in the field, only 665 of them were classified by Landsat 5 TM to be winter wheat, while 53 of them were classified as pasture and 100 were unclassified, for an accuracy of 81.3 percent.

Note: Urban/Quarry pixels were not subjected to ground truthing.

Table 3 Comparative Accuracy of In Situ Data Versus Image Data

		In Situ Data								
Landsat 5 TM Image Data		Forest	Pasture	Winter Wheat	Alfalfa	Corn	Soybean	Water	Total Pixels	Accuracy (%)
	Unclassified	2	-	100	-	2	1	2	107	N/A
	Forest	566	-	-	-	13	-	2	581	97.4
	Pasture	1	477	53	-	-	-	2	533	89.5
	Winter Wheat	-	-	665	-	-	-	-	665	100
	Alfalfa	-	21	-	58	-	-	-	79	73.4
	Corn	12	-	-	-	552	-	1	565	97.7
	Soybean	-	-	-	-	-	38	-	38	100
	Water	-	-	-	-	-	-	1259	1259	100
	Total Pixels	581	498	818	58	567	39	1266	3827	N/A
Accuracy (%)	97.4	95.8	81.3	100	97.3	97.4	99.4	N/A	Overall Accuracy = 89 %	

3 DATA ACQUISITION AND PROCESSING

3.1 Data Acquisition Methods

Oklahoma land use classification data were derived from both Landsat 5 TM imagery and in situ data, then classified and cross-verified for accuracy. These data were then compared to SGP97 and SGP99 land use classification data, two experiments which were conducted in the same region as SMEX03. Landsat 5 TM data were collected for 10 July 2003. Data were used from Path 28, Row 34-35 to cover the SMEX03 Oklahoma study area. In situ data were collected from 2 through 18 July 2003.

3.2 Derivation Techniques and Algorithms

3.2.1 Processing Steps

One scene for 10 July 2003 was used to construct the land use classification for the SMEX03 Oklahoma study area. There was little cloud cover for this particular date and other dates in proximity to the experiment were of low quality.

1. In situ data were collected in and around Chickasha, El Reno, and the Little Washita Basin in Oklahoma, USA.
2. In situ data were converted to Regions of Interest (ROIs) within ENVI and one third of these data were set aside for verification purposes.
3. Landsat TM data for 10 July 2003 were collected and imported to ENVI for Path 28/Row 34-35.
4. Landsat TM images were masked for clouds. The pixels with values in band 1 (.45 - .52 um) greater than 110 were removed.
5. Shadowed areas in Band 4 (.76 -.90um) with values less than 50 were removed.
6. Bands 2,3,4,5,7 were extracted from 10 July 2003 and merged to create a 5 band image.
7. Within ENVI, a supervised maximum likelihood classification was conducted with a threshold of .85.
8. A classification image was available from [SGP99](#). For areas that were masked due to clouds or shadows in the original classified image, SGP99 classifications were substituted, creating a merged image. Road network information was also superimposed on the image and those pixels selected as Urban/Quarry.
9. Accuracy percentages were calculated using a set of ROIs that were set aside for verification. Refer to Table 3 in the Quality Assessment section of this guide document for a listing of accuracy percentages.

4 REFERENCES AND RELATED PUBLICATIONS

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan. 2004. Development of a 2001 National Land-Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing* 70:829-840.

4.1 Related Websites

[The SGP99 Campaign](#)

[The Southern Great Plains '97 \(SGP97\) Data Archive](#)

5 CONTACTS AND ACKNOWLEDGMENTS

Thomas J. Jackson

Hydrology/Remote Sensing Laboratory

US Department of Agriculture (USDA) - Agricultural Research Service (ARS)

Bldg. 007, Rm. 104, BARC-West

Beltsville, Maryland 20705 USA

Michael H. Cosh

Hydrology/Remote Sensing Laboratory

US Department of Agriculture (USDA) - Agricultural Research Service (ARS)

Bldg. 007, Rm. 104, BARC-West

Beltsville, MD 20705 USA

Acknowledgments:

The investigators would like to acknowledge the [USDA ARS Grazinglands Research Laboratory](#) for their support as well as the many graduate students and volunteers who collected the field data.

6 DOCUMENT INFORMATION

6.1 Publication Date

January 2009

6.2 Date Last Updated

25 March 2021