

Применение дискриминантного анализа для распознавания посевов сельскохозяйственных культур

Э.А. Терехин

*Белгородский государственный национальный исследовательский университет
Федерально-региональный центр аэрокосмического и наземного мониторинга объектов
и природных ресурсов, Белгород, 308015, Россия
E-mail: terekhin@bsu.edu.ru*

Автоматизированное распознавание посевов является одной из ключевых задач мониторинга пахотных земель. В статье представлены результаты применения пошагового дискриминантного анализа для выявления посевов сельскохозяйственных культур, типичных для территории Белгородской области и Центрального Черноземья: пшеницы озимой, ячменя, сои, кукурузы, подсолнечника, сахарной свеклы, овса и многолетних трав. Экспериментальная информация получена с 1033 обрабатываемых полей. Спектральные характеристики сельскохозяйственной растительности рассчитаны на основе информационных продуктов MOD09Q1 и MOD13Q1. Эффективность выявления посевов проанализирована для рядов сезонных значений коэффициентов отражения в красной (620–670 нм) и ближней инфракрасной (841–876 нм) областях спектра и вегетационного индекса NDVI. Применение дискриминантного анализа позволило количественно оценить эффективность спектральных показателей конкретных временных срезов для выявления типов растительности. Установлено, что значения коэффициентов отражения либо вегетационного индекса середины – второй половины июля вносят наибольший вклад в разделение посевов. Наиболее высокая суммарная точность распознавания культур (около 85%) установлена для коэффициентов отражения в ближней инфракрасной зоне спектра. Наибольшая точность дешифрирования характерна для посевов озимых (96%). Распознавание многолетних трав и овса наиболее проблематично. Показана возможность применения дискриминантного анализа для геоинформационного картографирования растительности посевных площадей.

Ключевые слова: растительный покров, посевные площади, дискриминантный анализ, вегетационные индексы, коэффициенты отражения, NDVI, дистанционное зондирование

*Одобрена к печати: 19.07.2017
DOI: 10.21046/2070-7401-2017-14-7-89-99*

Введение

Проблема эффективного автоматизированного дешифрирования типов посевов и структуры севооборотов является одной из первоочередных задач при изучении растительности агроэкосистем. Ее решение обусловлено как научными, так и практическими целями. Для распознавания сельскохозяйственных культур могут быть использованы методы классификации с обучением и без него, однако они не всегда дают удовлетворительный результат. В ряде работ (Барталев и др., 2006; Atzberger, Rembold, 2013) показано, что в связи со сложностью и высокой динамичностью спектрально-отражательных свойств посевов для их эффективного выявления необходимо применение серий спутниковых изображений, полученных за вегетационный сезон. На их основе, в частности, предложены методы, применяющие локально-адаптивную классификацию (Барталев и др., 2011; Bartalev, Plotnikov, Loupian, 2016), специализированные вегетационные индексы (Плотников и др., 2011). Использование снимков высокого пространственного разрешения, как правило, позволяет получить более точные результаты при дешифрировании посевов (Повх, Шляхова, Боева, 2013). Но применение таких спутниковых данных ограничено их невысоким временным разрешением, что снижает возможности использования снимков для многих регионов с относительно большим количеством облачных дней в году, в частности для Белгородской области и Центрального Черноземья. Для распознавания посевов может быть использован попиксельный подход либо методы

зональной статистики, основанные на вычислении спектрально-отражательных характеристик в пределах контуров посевных площадей.

Распознавание посевов на основе серий спутниковых изображений обуславливает необходимость обработки достаточно большого количества независимых переменных, которыми являются значения спектральных показателей за конкретные временные срезы. Их эффективная оценка может быть проведена с применением подхода, использующего многомерный анализ данных. Возможность учета большого количества входных параметров обусловила все более широкое внедрение этой группы методов для решения географических задач (Кузьменко, Лисецкий, Пичура, 2012; Li et al., 2013; Zhou, Jia, Menenti, 2015). Для исследования состояния растительного покрова оценка возможности такого учета представляет особый интерес как в региональных, так и глобальных масштабах (Медведева и др., 2011; Лупян и др., 2014; Попов, 2016).

К настоящему времени накоплен архив ежедневно получаемых данных MODIS, позволяющий сформировать ряды сезонных значений спектральных вегетационных индексов или спектральных коэффициентов отражения, которые могут быть применены для изучения структуры севооборотов и анализа состояния растительности посевных площадей (Терехин, 2015; Son et al., 2014). В то же время оценка значительных массивов зональных коэффициентов отражения или вегетационных индексов, относящихся к различным типам посевов, требует использования методов многомерного анализа, одним из которых является дискриминантный анализ. Он является статистическим методом, позволяющим оценить разделение классов анализируемых объектов (в нашем случае типов посевов) на основе их определенных признаков (спектрально-отражательных характеристик).

Цель исследования заключалась в оценке возможностей дискриминантного анализа, основанного на совокупности сезонных значений спектральных показателей для изучения структуры севооборотов. Была поставлена задача определения эффективности сезонных коэффициентов отражения (в красной и ближней инфракрасной зоне спектра) и индекса NDVI для автоматизированного распознавания видов посевов, типичных для Белгородской области и Центрально-Черноземного региона.

Материалы и методика исследования

В исследовании было поставлено предположение, что совокупность сезонных значений вегетационного индекса либо коэффициентов отражения за весь период вегетации, рассчитанная для конкретных типов посевов, может быть использована для их выявления автоматизированным способом. Цель дискриминантного анализа заключается в расчете уравнений (функций), позволяющих определить наличие на конкретной посевной площади в текущем году того или иного типа посевов. Для каждого из них рассчитывается уравнение со своими коэффициентами. Исходя из того, что анализировалось восемь типов посевов (пшеница озимая, ячмень, соя, сахарная свекла, подсолнечник, овес, кукуруза и многолетние травы), было рассчитано по восемь функций для каждого (их трех) оцениваемого спектрального показателя.

Независимыми переменными для каждой функции выступали значения NDVI либо показания коэффициентов отражения в красной или ближней инфракрасной зоне спектра. Исследование было проведено в несколько этапов.

На первом из них был подготовлен векторный слой анализируемых посевных площадей и заполнена атрибутивная информация о типах посевов. Для создания векторного слоя были использованы снимки SPOT-5 (пространственное разрешение 2,5–5 м/пиксель), а для уточнения контуров полей на анализируемый 2012 г. — снимки Landsat TM, ETM+ (15–30 м/пиксель). Исходя из того, что в Белгородской области, как и на территории Центрального Черноземья, выращиваются различные типы сельскохозяйственной растительности, а севообороты характеризуются их сложным сочетанием, исследование требовало подборки значительной по объему выборки посевных площадей с различными типами культур, произрастающих на них. Для этого нами на основе данных сельскохозяйственных организаций и собственных полевых исследований был выполнен сбор информации с 1033-х посевных площадей (*табл. 1*), расположенных в различных районах области, содержащей сведения о типах посевах за 2012 г. Таким образом, в исследовании был охвачен практически полный набор культур, составляющих основу севооборотов Белгородской области. Общая площадь анализируемых обрабатываемых полей составила свыше 64 000 га.

Второй этап заключался в обработке информации о сезонной динамике спектрально-отражательных свойств посевных площадей, которая была получена на основе информационных продуктов MOD09Q1 и MOD13Q1 (<https://lpdaac.usgs.gov/>). Изображения MOD09Q1 содержат 8-дневные композитные изображения коэффициентов отражения в красной (620–670 нм) и ближней инфракрасной (841–876 нм) областях спектра. Они рассчитаны на основе атмосферно и радиометрически откорректированных спутниковых данных MODIS (пространственное разрешение 250 м/пиксель).

Таблица 1. Количество проанализированных полей под различными типами посевов

<i>Тип посевов</i>	<i>Число анализируемых полей</i>	<i>Средняя площадь, га</i>	<i>Суммарная площадь, га</i>
Озимая пшеница	372	64,4	23 957,5
Соя	262	64,7	16 943,2
Сахарная свекла	103	71,8	7 395,5
Подсолнечник	84	59,2	4 976,2
Ячмень	23	67,3	1 547,4
Овес	42	48,8	2 047,9
Многолетние травы	41	28,6	1 174,5
Кукуруза	106	57,4	6 087,7
Всего	1033	62,1	64 129,9

Продукты MOD13Q1 являются 16-дневными композитными изображениями вегетационного индекса NDVI (нормализованный разностный вегетационный индекс) за интервал в 16 дней. Стоит отметить, что на текущий момент времени снимки MODIS являются единственным типом спутниковых данных, позволяющим благодаря высокому временному (и достаточному пространственному) разрешению провести планируемое исследование.

Для каждой анализируемой посевной площади на основе изображений MOD13Q1 был рассчитан сезонный ряд вегетационного индекса NDVI за вегетационный период 2012 г. (по 15 временных срезов). После этого аналогично для каждого обрабатываемого поля на основе данных MOD09Q1 рассчитывался сезонный ряд коэффициентов отражения в красной и ближней инфракрасной зоне спектра (по 27 временных срезов). По причине наличия облачности изображения MOD09Q1 некоторых дат 2012 г. пришлось исключить из анализа (24.05, 01.06, 12.08, 28.08 и 23.10).

Значения вегетационного индекса и коэффициентов отражения для конкретных обрабатываемых полей были рассчитаны в программе ERDAS IMAGINE методом зональной статистики, т.е. на основе растров, содержащих значения спектральных характеристик, вычислялись их средние значения в пределах контуров посевных площадей. Применение такого подхода позволило сделать анализируемым объектом конкретное сельскохозяйственное поле, но, с другой стороны, относительно невысокое пространственное разрешение исходных данных (250 м) снизило достоверность расчета спектральных показателей для малых по размеру полей (менее 15–20 га). Для удобства анализа были использованы значения вегетационного индекса и коэффициентов отражения (являющихся безразмерной величиной), умноженные на 10 000. В анализе были использованы значения NDVI и коэффициентов отражения с начала марта по начало ноября. Таким образом, динамика зональных коэффициентов отражения была изучена с интервалом в 8 дней, вегетационного индекса NDVI — с интервалом в 16 дней. Использование осенних значений спектрально-отражательных характеристик обусловлено тем, что некоторые типы посевов, выращиваемые в регионе (сахарная свекла, многолетние травы), находятся на полях до конца октября.

Этап статистической обработки данных заключался в их оценке с применением методов дискриминантного анализа. Каждой исследуемой посевной площади был присвоен уникальный номер, позволивший провести сопоставление результатов в геоинформационной системе ArcGIS и статистическом пакете STATISTICA. Нами был использован метод пошагового дискриминантного анализа (Халафян, 2007), позволяющий отобрать по величине F-критерия независимые переменные (в нашем случае — спектральные показатели), вносящие наибольший вклад в разделение посевов, и впоследствии использовать их в дискриминантных функциях. Анализ был проведен отдельно для каждой спектральной характеристики (рис. 1).

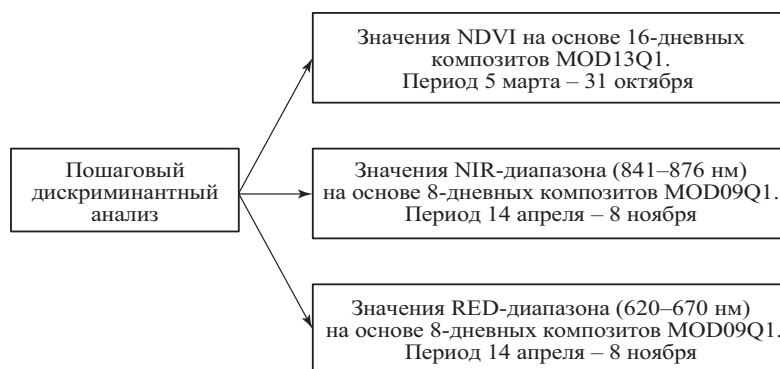


Рис. 1. Независимые переменные (спектральные показатели), эффективность которых была изучена для распознавания типов посевов

Обучающую выборку для построения уравнений классификации проводили на основе данных с 450 случайно отобранных посевных площадей (из 1033), а затем проверяли эффективность рассчитанных дискриминантных функций на информации с остальных 583 полей.

Результаты и их обсуждение

Пошаговый дискриминантный анализ позволил подробно проанализировать эффективность сезонных значений NDVI, коэффициентов отражения (коэффициентов спектральной яркости, КСЯ) в красном и ближнем инфракрасном диапазонах для разделения типов посевов (*табл. 2*). Переменные за каждую дату при выполнении процедуры последовательно включались в модель по степени их вклада в дискриминацию посевов (по наибольшей величине F-критерия).

Таблица 2. Результаты пошагового дискриминантного анализа, характеризующие эффективность сезонных значений спектральных показателей 2012 года для разделения посевов

<i>КСЯ RED-диапазона</i>		<i>КСЯ NIR-диапазона</i>		<i>NDVI</i>	
<i>Дата</i>	<i>F-включение</i>	<i>Дата</i>	<i>F-включение</i>	<i>Дата</i>	<i>F-включение</i>
11.07	108,27	19.07	118,51	11.07	155,72
16.05	24,23	04.08	40,27	29.09	32,54
04.08	20,58	30.04	39,17	22.04	31,31
14.04	18,46	16.05	26,77	12.08	23,38
21.09	11,49	08.11	19,05	08.05	15,86
09.06	11,25	29.09	12,92	13.09	10,51
15.10	10,60	03.07	11,76	31.10	6,74
13.09	10,46	09.06	8,33	15.10	6,20
17.06	6,60	15.10	7,12	05.03	6,00
22.04	6,47	21.09	5,10	28.08	3,71
08.11	6,19	14.04	4,66		
27.07	5,80	08.05	4,18		
03.07	4,59	25.06	3,58		
25.06	3,26	27.07	3,12		
30.04	3,81				

В *табл. 2* представлены только статистически значимые переменные (используемые впоследствии для расчета коэффициентов дискриминантных функций), большие по величине порогового значения критерия F. Из таблицы видно, что наибольший вклад в разделение посевов вносят показания спектральных характеристик середины – второй половины июля, т.е. можно сделать вывод о высокой эффективности значений NDVI либо зональных коэффициентов отражения этого периода для дешифрирования посевов.

Для остальных дат каждый спектральный показатель имеет свой набор эффективных характеристик. Переменные, представленные в *табл. 2*, были использованы для расчета коэффициентов дискриминантных функций, т.е. уравнения, основанные на сезонных значениях

RED-диапазона, имели 15 переменных, NIR-диапазона — 14 переменных, NDVI — 10 переменных.

Оценка точности распознавания посевов (табл. 3–5) на основе рассчитанных дискриминантных функций показала, что суммарная точность классификации изменяется от 81,5% (NDVI) до почти 85,0% (КСЯ NIR-диапазона). При этом необходимо отметить, что для расчета зональных коэффициентов отражения использовались 8-дневные композиты, а для NDVI — 16-дневные. Более высокая точность дешифрирования посевов по значениям NIR-диапазона также может быть обусловлена тем, что отношение сигнал/шум в красном канале MODIS значительно хуже, чем в инфракрасном. Максимальная точность дешифрирования отдельных типов культур, например пшеницы озимой, достигает 95%.

Таблица 3. Точность распознавания посевов на основе дискриминантного анализа коэффициентов отражения в ближнем инфракрасном (NIR) диапазоне, 841–876 нм

Тип посевов	Точность распознавания, %	Ячмень	Соя	Сахарная свекла	Подсолнечник	Озимая пшеница	Овес	Многолетние травы	Кукуруза
Ячмень	72,7	8	0	0	0	3	0	0	0
Соя	81,3	0	65	4	1	3	1	0	6
Сахарная свекла	61,1	0	5	22	0	1	0	1	7
Подсолнечник	72,2	0	3	0	26	6	0	0	1
Озимая пшеница	95,6	0	1	0	0	172	2	3	2
Овес	66,7	0	0	0	0	4	8	0	0
Многолетние травы	66,7	0	1	0	0	4	0	10	0
Кукуруза	83,9	0	2	2	0	5	0	0	47
Всего	84,8	8	77	28	27	198	11	14	63

Из проанализированных спектральных показателей наиболее высокую суммарную точность распознавания посевов показали коэффициенты отражения в ближней инфракрасной области спектра (841–876 нм), рассчитанные на основе 8-дневных композитов. Наибольшая точность распознавания характерна для пшеницы озимой, наименьшая — для многолетних трав и овса. Более высокая точность распознавания большинства типов посевов на основе данных NIR-диапазона по сравнению с RED-диапазоном, скорее всего, обусловлена большими различиями в спектрально-отражательных свойствах растительности в этой спектральной зоне. Также она может быть обусловлена тем, что сами значения яркости в красном канале больше подвержены влиянию состояния атмосферы, несмотря на атмосферную коррекцию и осреднение по времени. По сравнению с NDVI она может являться следствием применения 8-дневных композитов по сравнению с 16-дневными, используемыми для расчета значений вегетационного индекса.

Таблица 4. Точность распознавания посевов на основе дискриминантного анализа коэффициентов отражения в красном (RED) диапазоне, 620–670 нм

<i>Тип посевов</i>	<i>Точность распознавания, %</i>	<i>Ячмень</i>	<i>Соя</i>	<i>Сахарная свекла</i>	<i>Подсолнечник</i>	<i>Озимая пшеница</i>	<i>Овес</i>	<i>Многолетние травы</i>	<i>Кукуруза</i>
Ячмень	81,8	9	0	0	0	2	0	0	0
Соя	76,3	0	61	5	3	3	1	0	7
Сахарная свекла	58,3	0	7	21	1	0	0	1	6
Подсолнечник	72,2	0	5	1	26	3	0	0	1
Озимая пшеница	95,0	2	2	0	1	171	0	2	2
Овес	75,0	0	0	2	0	1	9	0	0
Многолетние травы	46,7	0	1	1	1	5	0	7	0
Кукуруза	76,8	0	5	2	4	1	0	1	43
Всего	81,5	11	81	32	36	186	10	11	59

Таблица 5. Точность распознавания посевов на основе дискриминантного анализа значений вегетационного индекса NDVI

<i>Тип посевов</i>	<i>Точность распознавания, %</i>	<i>Ячмень</i>	<i>Соя</i>	<i>Сахарная свекла</i>	<i>Подсолнечник</i>	<i>Озимая пшеница</i>	<i>Овес</i>	<i>Многолетние травы</i>	<i>Кукуруза</i>
Ячмень	90,9	10	0	0	0	0	1	0	0
Соя	73,8	0	59	1	7	2	0	0	11
Сахарная свекла	75,0	0	3	27	0	0	0	1	5
Подсолнечник	55,6	0	6	0	20	1	0	0	9
Озимая пшеница	95,0	1	2	0	0	171	2	2	2
Овес	58,3	2	1	1	0	1	7	0	0
Многолетние травы	66,7	0	0	0	0	4	1	10	0
Кукуруза	76,8	0	7	3	1	1	0	1	43
Всего	81,5	13	78	32	28	180	11	14	70

Примерно одинаковая суммарная точность распознавания культур (81,5%) на основе NDVI и коэффициентов отражения в красном диапазоне (с учетом того, что для расчета коэффициентов отражения были использованы 8-дневные, а не 16-дневные композиты) указывает на более высокую эффективность значений вегетационного индекса для распознавания

посевов. Кроме того, наименьшую эффективность коэффициентов отражения в красной зоне спектра по сравнению с остальными спектральными показателями обуславливает минимальная выявленная (46,7%) точность распознавания посевов, наблюдающаяся для многолетних трав.

Полученные результаты позволили выполнить экспериментальное геоинформационное картографирование структуры севооборотов (на основе NIR-диапазона), пример которого представлен на *рис. 2*.

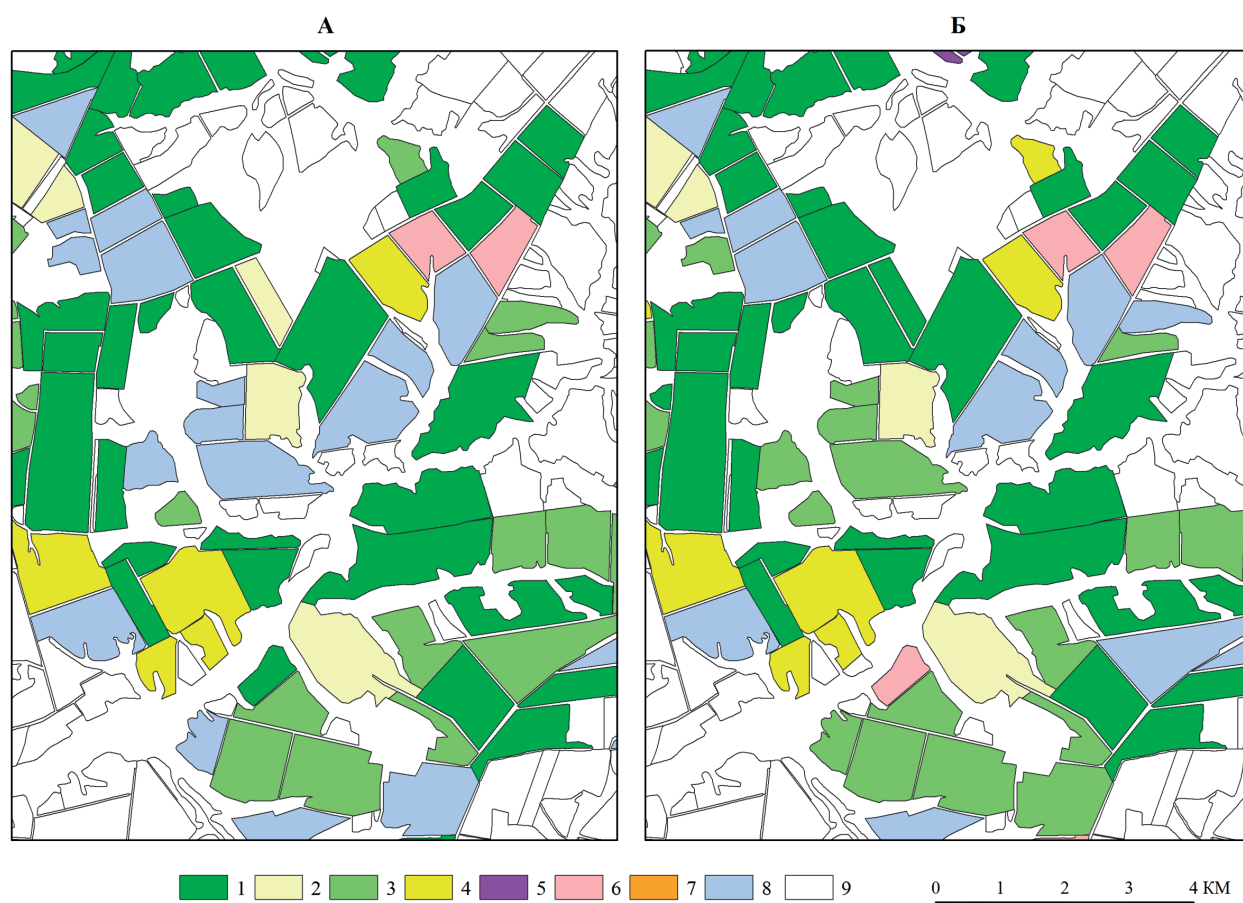


Рис. 2. Пример сравнения реального расположения посевов (А) с результатами их распознавания (Б), полученными с применением дискриминантного анализа: 1 — пшеница озимая; 2 — ячмень; 3 — соя; 4 — кукуруза; 5 — многолетние травы; 6 — овес; 7 — подсолнечник; 8 — сахарная свекла; 9 — не анализируемые посевные площади

Таким образом, наибольшую эффективность для автоматизированного распознавания посевов показали значения коэффициентов отражения в ближней инфракрасной зоне спектра, рассчитанные на основе 8-дневных композитов. Максимальная точность распознавания культур на их основе составила 95,6% для пшеницы озимой, минимальная — для овса и многолетних трав — 66,7%. Относительно невысокая точность распознавания этих типов посевов обусловлена тем, что они характеризуются близкими спектрально-отражательными свойствами.

Главное преимущество в применении дискриминантного анализа заключается в возможности учета большого количества независимых переменных (в нашем случае — от 15

до 27) при распознавании посевов и отборе наиболее информативных сезонных значений спектральных характеристик. Основным недостатком используемого подхода обусловлен снижением достоверности расчета спектральных показателей для небольших по размеру посевных площадей, причиной чего является невысокое пространственное разрешение исходных данных.

Выводы

С применением пошагового дискриминантного анализа изучена эффективность коэффициентов отражения в красном (620–670 нм) и ближнем инфракрасном (841–876 нм) диапазонах и вегетационного индекса NDVI для автоматизированного распознавания посевов, типичных для Центрального Черноземья. Анализ экспериментальной выборки с 1033-х посевных площадей позволил установить, что значения всех оцениваемых спектральных показателей середины – второй половины июля вносят наибольший вклад в разделимость посевов. Для NDVI на следующем месте по эффективности следуют значения конца апреля и первой половины октября. Суммарная точность распознавания посевов (озимой пшеницы, ячменя, сои, кукурузы, овса, многолетних трав, подсолнечника и сахарной свеклы) составила около 85% для коэффициентов отражения в ближней инфракрасной спектральной зоне и по 81,5% для коэффициентов отражения в красной зоне спектра и значений вегетационного индекса. Максимальная точность распознавания составила около 96% для озимой пшеницы на основе коэффициентов отражения NIR-диапазона. Наиболее проблематично разделение посевов многолетних трав и овса. Снижение точности распознавания некоторых культур обусловлено схожестью их спектрально-отражательных свойств и наличием небольших по размеру посевных площадей.

Литература

1. *Барталев С.А., Лупян Е.А., Нейштадт И.А., Савин И.Ю.* Классификация некоторых типов сельскохозяйственных посевов в южных регионах России по спутниковым данным MODIS // Исследование Земли из космоса. 2006. № 3. С. 68–75.
2. *Барталев С.А., Егоров В.А., Лупян Е.А., Плотников Д.Е., Уваров И.А.* Распознавание пахотных земель на основе многолетних спутниковых данных спектро радиометра MODIS и локально-адаптивной классификации // Компьютерная оптика. 2011. Т. 35. № 1. С. 103–116.
3. *Кузьменко Я.В., Лисецкий Ф.Н., Пичура В.И.* Оценка и прогнозирование стока малых рек в условиях антропогенных воздействий и изменений климата // Современные проблемы науки и образования. 2012. № 6. С. 619–627.
4. *Лупян Е.А., Барталев С.А., Толпин В.А., Жарко В.О., Крашенинникова Ю.С., Оксюкевич А.Ю.* Использование спутникового сервиса ВЕГА в региональных системах дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2014. Т. 11. № 3. С. 215–232.
5. *Медведева М.А., Савин И.Ю., Барталев С.А., Лупян Е.А.* Использование данных NOAA-AVHRR для выявления многолетней динамики растительности северной Евразии // Исследование Земли из космоса. 2011. № 4. С. 55–62.
6. *Плотников Д.Е., Барталев С.А., Жарко В.О., Михайлов В.В., Присянникова О.И.* Экспериментальная оценка распознаваемости агрокультур по данным сезонных спутниковых измерений спектральной яркости // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 199–208.
7. *Повх В.И., Шляхова Л.А., Боева И.Н.* Оценка структуры посевных площадей сельскохозяйственных культур по спутниковой информации высокого разрешения // Современные проблемы дистанционного зондирования Земли из космоса. 2013. Т. 10. № 4. С. 224–228.
8. *Попов С.Ю.* Опыт создания геоботанической карты методом дискриминантного анализа полевых и дистанционных данных // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13. № 1. С. 25–35.

9. Терехин Э.А. Анализ многолетней динамики вегетационного индекса для посевных площадей // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 6. С. 48–58.
10. Халафян А.А. STATISTICA 6. Статистический анализ данных. М.: Бином-Пресс. 2007. 512 с.
11. Atzberger C., Rembold F. Mapping the spatial distribution of winter crops at sub-pixel level using AVHRR NDVI time series and neural nets // Remote Sensing. 2013. Vol. 5. No. 3. P. 1335–1354.
12. Bartalev S.A., Plotnikov D.E., Loupian E.A. Mapping of arable land in Russia using multi-year time series of MODIS data and the LAGMA classification technique // Remote Sensing Letters. 2016. Vol. 7. No. 3. P. 269–278.
13. Li Z., Huffman T., McConkey B., Townley-Smith L. Monitoring and modeling spatial and temporal patterns of grassland dynamics using time-series MODIS NDVI with climate and stocking data // Remote Sensing of Environment. 2013. Vol. 138. P. 232–244.
14. Son N.T., Chen C.F., Chen C.R., Minh V.Q., Trung N.H. A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. Agricultural and Forest Meteorology. 2014. Vol. 197. P. 52–64.
15. Zhou J., Jia L., Menenti M. Reconstruction of global MODIS NDVI time series: performance of harmonic analysis of time series (hants) // Remote Sensing of Environment. 2015. Vol. 163. P. 217–228.

Application of discriminant analysis for recognition of agricultural crops

E.A. Terekhin

Belgorod State National Research University, Belgorod 308015, Russia
E-mail: terekhin@bsu.edu.ru

Automated recognition of agricultural crops is one of key problems in monitoring of arable lands. This paper presents the results of stepwise discriminant analysis employed to recognize crops typical of the Belgorod Oblast and Central Black Earth Region: winter wheat, barley, soybean, corn, sunflower, sugar beet, oat and perennial grasses. Experimental data were received from 1033 sowing areas. Spectral parameters of crops were calculated using MOD09Q1 and MOD13Q1 data. Efficiency of crops recognition was analyzed for the time series of reflectance in the red (620-670 nm) and near infrared (841-876 nm) bands and NDVI vegetation index. The use of discriminant analysis made it possible to assess the utility of the seasonal values of spectral parameters for identification of crops types. The values of all spectral reflection parameters of the middle of July or its second half contribute the most to the distinction of the crops. The highest total crops recognition accuracy (about 85%) was established for near infrared reflectance (841-876 nm). The highest accuracy of interpretation was established for winter cereals (96%). The recognition of perennial grasses and oats is the most problematic. The possibility of using discriminant analysis in geoinformation mapping of the crops was shown.

Keywords: vegetation cover, sown areas, stepwise discriminant analysis, vegetation indices, spectral reflectance, NDVI, remote sensing

Accepted: 19.07.2017

DOI: 10.21046/2070-7401-2017-14-7-89-99

References

1. Bartalev S.A., Lupyan E.A., Neishtadt I.A., Savin I.Yu., Klassifikatsiya nekotorykh tipov sel'skokhozyaystvennykh posevov v yuzhnykh regionakh Rossii po sputnikovym dannym MODIS (Gropland Area Classification in South Regions of Russia Using MODIS Satellite Data), *Issledovanie Zemli iz kosmosa*, 2006, No. 3, pp. 68–75.
2. Bartalev S.A., Egorov V.A., Loupian E.A., Plotnikov D.E., Uvarov I.A., Raspoznavanie pakhotnykh zemel' na osnove mnogoletnikh sputnikovyykh dannyykh spektroradiometra MODIS i lokal'no-adaptivnoi klassifikatsii (Recognition of arable lands using multi-annual satellite data from spectroradiometer MODIS and locally adaptive supervised classification), *Komp'yuternaya optika*, 2011, Vol. 35. No. 1, pp. 103–116.
3. Kuz'menko Ya.V., Lisetskii F.N., Pichura V.I., Otsenka i prognozirovanie stoka malykh rek v usloviyakh antropogennykh vozdeystvii i izmenenii klimata (Evaluation and prediction of small rivers flow in condition of anthropogenic impacts and climate change), *Sovremennye problemy nauki i obrazovaniya*, 2012, No. 6, pp. 619–627.
4. Loupian E.A., Bartalev S.A., Tolpin V.A., Zharko V.O., Krashennnikova Yu.S., Oksyukevich A.Yu., Ispol'zovanie sputnikovogo servisa VEGA v regional'nykh sistemakh distantsionnogo monitoring (VEGA satellite service

- applications in regional remote monitoring systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2014, Vol. 11, No. 3, pp. 215–232.
5. Medvedeva M.A., Savin I.Yu., Bartalev S.A., Loupian E.A., Ispol'zovanie dannykh NOAA-AVHRR dlya vyyavleniya mnogoletnei dinamiki rastitel'nosti severnoi Evrazii (Usage of NOAA-AVHRR Data for Detection of Vegetation Dynamics in Northern Eurasia), *Issledovanie Zemli iz kosmosa*, 2011, No. 4, pp. 55–62.
 6. Plotnikov D.E., Bartalev S.A., Zharko V.O., Mikhailov V.V., Prosyannikova O.I., Eksperimental'naya otsenka raspoznavаемosti agrokul'tur po dannym sezonnykh sputnikovyykh izmerenii spektral'noi yarkosti (An experimental assessment of crop types recognisability using time-series of intra-seasonal spectral reflectance measurements by satellite sensor), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 1, pp. 199–208.
 7. Povkh V.I., Shlyakhova L.A., Boeva I.N., Otsenka struktury posevnykh ploshchadei sel'skokhozyaistvennykh kul'tur po sputnikovoi informatsii vysokogo razresheniya (Assessment of the structure of agricultural crop areas using high resolution satellite data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2013, Vol. 10, No. 4, pp. 224–228.
 8. Popov S.Yu. Opyt sozdaniya geobotanicheskoi karty metodom diskriminantnogo analiza polevykh i distantsionnykh dannykh (Experience of creating a geobotanical map using discriminant analysis of field vegetation description and remote sensing data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2016, Vol. 1, No. 1, pp. 25–35.
 9. Terekhin E.A., Analiz mnogoletnei dinamiki vegetatsionnogo indeksa dlya posevnykh ploshchadei (Analysis of vegetation index long-term dynamics for crop areas), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 6, pp. 48–58.
 10. Khalafyan A.A. *STATISTICA 6. Statisticheskii analiz dannykh* (STATISTICA 6. Statistical analysis), Moscow: Binom-Press, 2007, 512 p.
 11. Atzberger C., Rembold F., Mapping the spatial distribution of winter crops at sub-pixel level using AVHRR NDVI time series and neural nets, *Remote Sensing*, 2013, Vol. 5, No. 3, pp. 1335–1354.
 12. Bartalev S.A., Plotnikov D.E., Loupian E.A., Mapping of arable land in Russia using multi-year time series of MODIS data and the LAGMA classification technique, *Remote Sensing Letters*, 2016, Vol. 7, No. 3, pp. 269–278.
 13. Li Z., Huffman T., McConkey B., Townley-Smith L., Monitoring and modeling spatial and temporal patterns of grassland dynamics using time-series MODIS NDVI with climate and stocking data, *Remote Sensing of Environment*, 2013, Vol. 138, pp. 232–244.
 14. Son N.T., Chen C.F., Chen C.R., Minh V.Q., Trung N.H., A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation, *Agricultural and Forest Meteorology*, 2014, Vol. 197, pp. 52–64.
 15. Zhou J., Jia L., Menenti M., Reconstruction of global MODIS NDVI time series: performance of harmonic analysis of time series (hants), *Remote Sensing of Environment*, 2015, Vol. 163, pp. 217–228.