

Информатика, вычислительная техника и управление

УДК 004.932

АНАЛИЗ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ ДЛЯ ОБНАРУЖЕНИЯ НЕФТЯНЫХ РАЗЛИВОВ

В.А. Беликов, В.В. Галянин, С.П. Орлов

Самарский государственный технический университет
Россия, 443100, г. Самара, ул. Молодогвардейская, 244

Рассматривается проблема обнаружения нефтяных разливов, вызванных производственной деятельностью человека. Предложено использовать web-сервис хемометрических инструментов для построения PCA, PLS-DA моделей для выявления нефтяных компонентов на поверхности. Рассмотрены примеры построенных моделей для обнаружения нефтепродуктов на земной поверхности в районах Самарской области. Показано, что получаемая сепарация нефтепродуктов, водных ресурсов, городской застройки, растительности и почвы имеет высокую точность. С помощью перекрестной проверки произведена количественная оценка погрешности используемых моделей.

Ключевые слова: моделирование, web-сервис, дистанционное зондирование Земли, кластерный анализ, регрессионный анализ.

Введение

Нефтяные разливы наносят существенный ущерб экологии, они могут произойти на любой из стадий нефтедобычи, переработки или хранения. По данным Министерства природных ресурсов и экологии Российской Федерации, ежегодно в стране происходит свыше 25 тысяч аварийных нефтяных разливов, при этом на месторождениях вследствие аварийных разливов в окружающую среду поступает не менее 50 тысяч тон нефти и пластовых вод [1]. При этом до надзорных органов доходит информация лишь о небольшой их части. Причиной такого большого количества экологических происшествий чаще всего являются ветхие нефтепроводы, незаконные врезки и аварии на транспорте [1].

Сложно переоценить вред, наносимый экологии страны нефтяными разливами. Как правило, нефтяные разливы происходят в связи с прорывами ветхих трубопроводов, но произойти они могут на любой из стадий нефтедобычи, переработки или хранения. Их своевременное выявление и ликвидация снижают непоправимый ущерб экологии региона.

*Вадим Александрович Беликов, аспирант кафедры «Вычислительная техника».
Владислав Владимирович Галянин, заведующий лабораторией «Многомерный анализ и глобальное моделирование» СамГТУ.*

Сергей Павлович Орлов (д.т.н., проф.), заведующий кафедрой «Вычислительная техника».

На данный момент уже существуют методики обнаружения нефтяных разливов с помощью средств ДЗЗ (дистанционного зондирования Земли), однако они в большинстве своем направлены на детектирование нефти на открытой воде (заливы, моря, океаны) радарными измерениями либо на мониторинг какого-либо определенного протяженного участка поверхности Земли с помощью средств беспилотных летательных аппаратов [2, 3].

В статье описывается подход к определению нефтяных разливов на значительных территориях (в масштабе городов и областей) с использованием мультиспектральных данных с космического аппарата (КА) Landsat-8, методов хемометрики и машинного обучения [4, 5].

Космические миссии ДЗЗ

В настоящее время существует несколько глобальных космических миссий ДЗЗ, поставляющих результаты своей работы как на коммерческих, так и на свободных, некоммерческих условиях [2, 3].

Одна из таких миссий – программа USGS Landsat, которая использует серию одноименных космических аппаратов, оборудованных для получения мультиспектральных снимков. Последний из запущенных КА – Landsat-8, оснащенный сенсорами TIRS (Thermal InfraRed Sensor) и OLI (Operational Land Imager), позволяет получать мультиспектральные изображения поверхности Земли с разрешением 30 м для диапазонов 430–880; 1600; 2200 нм и приведенные к 30 м в диапазоне 10,6–12,5 мкм [6].

Аналогичную миссию ДЗЗ ведет Европейское космическое агентство – Copernicus. Пока на орбиту запущен один КА – Sentinel-2A, его оборудование позволяет получать мультиспектральные снимки с разрешением 20 метров для диапазонов 490–865; 1610; 2190 нм [7].

В России разработана и выведена в космос группировка КА Ресурс-П, оборудование которой позволяет получать гиперспектральные изображения с разрешением 30 м в диапазоне 400–1100 нм [8]. Система дистанционного зондирования Земли «Ресурс-П» состоит из гиперспектральной аппаратуры, комплекса широкозахватной мультиспектральной аппаратуры высокого и среднего разрешения, оптико-электронного комплекса («Геотон-Л1» с СППИ «Сангур-1У») и ряда вспомогательных устройств, обеспечивающих координацию, а также стабилизацию полета и съемки.

Снимки поверхности Земли, полученные миссиями Landsat и Copernicus, можно получить с сайтов этих проектов [6, 7].

Используемые изображения

В данной работе были использованы два изображения поверхности Земли, снятые космическим аппаратом Landsat-8. На первой фотографии снята территория г. Самары и области, на второй – г. Салавата Республики Башкортостан. Каждый такой снимок поверхности Земли охватывает территорию протяженностью 170 км с севера на юг и 185 км с востока на запад [6]. Оба снимка были получены в летние месяцы (июнь для Самары и август для Салавата) в условиях малой облачности и схожих температур воздуха.

Данные были скачаны с web-сервиса EarthExplorer Геологической службы США. Изображения, находящиеся в коллекции данных EarthExplorer, не являются полностью сырыми данными и были обработаны до уровня 1T, т. е. рельеф местности был скорректирован. Данный уровень коррекции подразумевает, что

рельеф местности будет максимально контрастен и точен. Также после коррекции пространственное разрешение изображений, полученных с каналов 10 и 11, изменяется со 100 до 30 м [6].

Пространственное разрешение каналов сенсоров космического аппарата Landsat-8 представлено в таблице.

Спектральные каналы космического аппарата Landsat-8

Спектральный канал	Длины волн, мкм	Пространственное разрешение, м
Канал 1 – Побережья и аэрозоли	0,433–0,453	30
Канал 2 – Синий	0,450–0,515	30
Канал 3 – Зеленый	0,525–0,600	30
Канал 4 – Красный	0,630–0,680	30
Канал 5 – Ближний ИК	0,845–0,885	30
Канал 6 – Ближний ИК	1,560–1,660	30
Канал 7 – Ближний ИК	2,100–2,300	30
Канал 8 – Панхроматический	0,500–0,680	15
Канал 9 – Перистые облака	1,360–1,390	30
Канал 10 – Дальний ИК	10,30–11,30	100 (30)
Канал 11 – Дальний ИК	11,50–12,50	100 (30)

Такое разрешение накладывает ограничения на возможность обнаружения нефтяных разливов, т. к. каждый пиксел отображает квадрат поверхности Земли со сторонами в 30 м. Таким образом, используя Landsat-8, можно обнаруживать только достаточно крупные нефтяные разливы.

Данные поставляются как отдельные изображения в формате GeoTIFF для каждого канала сенсоров спутника, что позволяет выборочно подходить к анализу. В данной работе используются 1–7 и 9–11 каналы, обладающие одинаковым пространственным разрешением в 30 м на 1 пиксел.

Моделирование

Для исследования возможности определения нефтяных разливов на суше с использованием КА программы Landsat были проанализированы изображения Самарской области, полученные со спутника Landsat-8.

Пиксели изображений были поделены на 4 класса: городская застройка, поле, вода и нефтешламы. В качестве эталонного участка класса «нефтешламы» было выбрано озеро Нефтешламовое (53.085005, 50.027618). С помощью сервиса TPT-cloud [9, 10] была построена PCA-модель (метод главных компонент) [4] и выявлены наиболее эффективные методы нормировки данных.

На графике счетов (рис. 1) хорошо видно сепарацию нефтешламов от трех других классов. Для дискриминации нефтепродуктов на поверхности Земли была построена регрессионная модель методом PLS-DA (метод проекций на латентные структуры для дискриминационного анализа) [5]. Результаты представлены на рис. 2.

Для проверки данного подхода в масштабе различных городов и областей была построена аналогичная модель одновременно по нескольким изображениям г. Самары и г. Салавата Республики Башкортостан. В качестве обучающего клас-

са «нефть» были взяты пиксели открытых резервуаров с нефтепродуктами (г. Салават) и озера Нефтьшамовое (г. Самара).

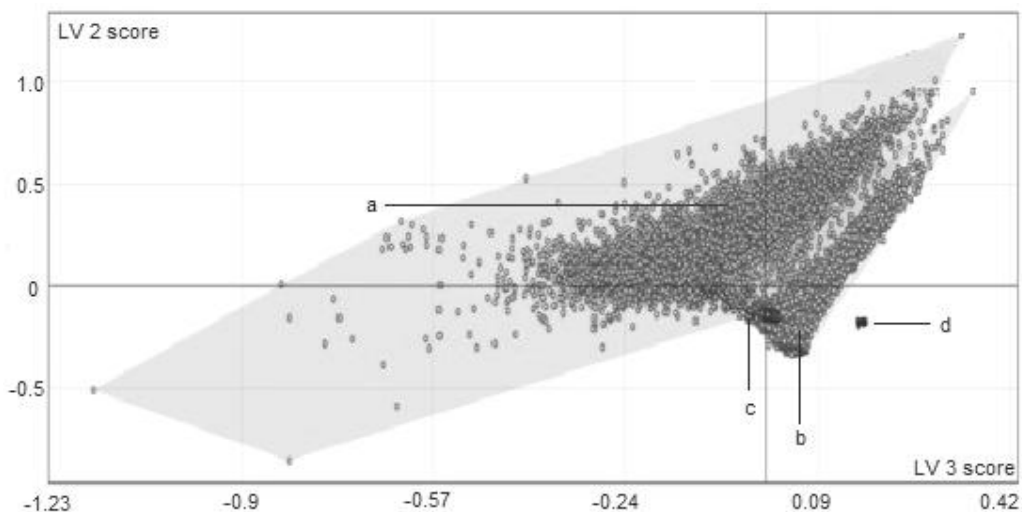


Рис. 1. Графиков счетов PCA-модели, построенной на мультиспектральных данных с КА Landsat-8 г. Снимок г. Самары, июнь 2014 г.:
а) городская застройка; б) поле; в) вода; д) нефтьшамы.

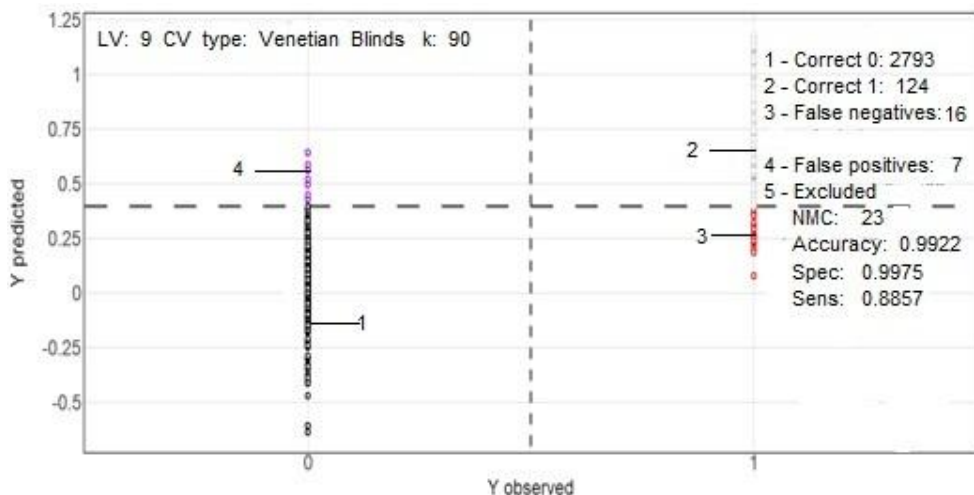


Рис. 2. График зависимости Y предсказанного от Y калибровочного модели PLS-DA, построенной на данных мультиспектрального изображения Landsat-8 г. Самары

В качестве предобработок были использованы методы нормировки, выбранные при PCA-моделировании. Полученная модель показала хорошее разделение нефтепродуктов, водных ресурсов, растительности и открытых участков почвы (рис. 3, 4).

На графике счетов (рис. 3) наблюдается хорошее отделение классов воды и нефтепродуктов от классов почв и растительности. При построении модели также использовались данные небольших водоемов рядом с нефтехранилищами – они обозначены как класс «вода – нефть» на графике. Это группа частично загрязнения, находящаяся между классами воды и нефтепродуктов.

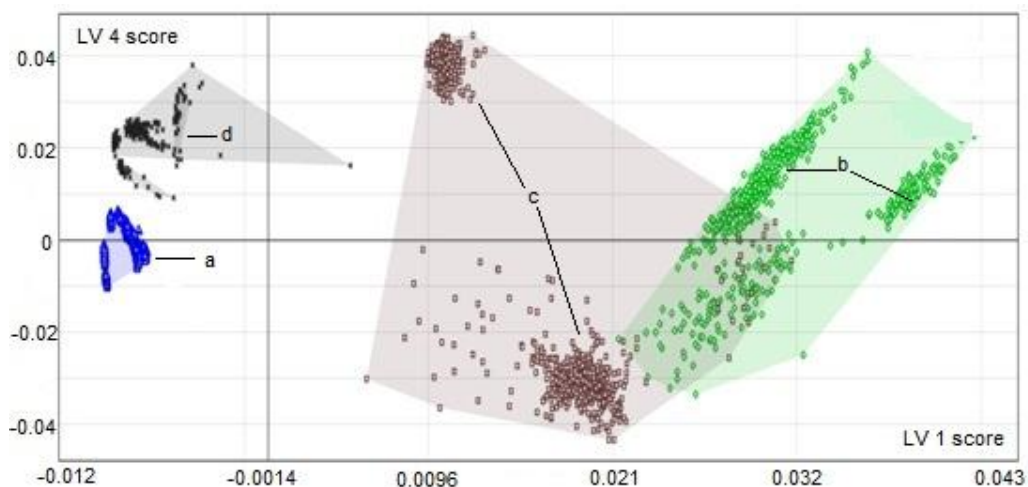


Рис. 3. График счетов PLS-DA модели, построенной на изображении г. Самары и г. Салавата:

а) вода; б) растительность; в) почва; д) нефтешламы

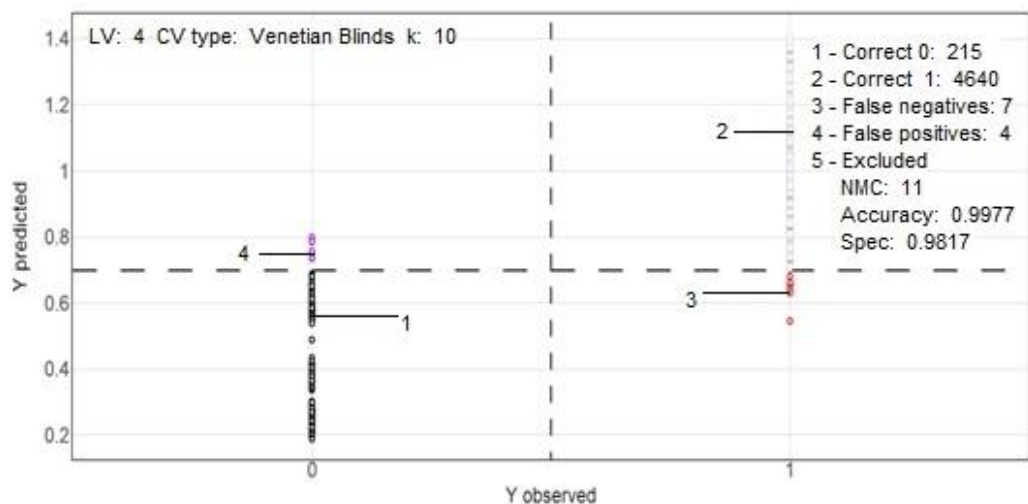


Рис. 4. График зависимости Y предсказанного от Y калибровочного модели PLS-DA, построенной на снимках г. Самары и г. Салавата

Для валидации получившейся модели была использована перекрестная проверка (ПП) с разделением калибровочного набора на 10 частей. Чувствительность модели (ПП) к объединенному классу «нефть + нефтешламы + воданефть» составила 99,85 %, а специфичность (ПП) 98,17 % (см. рис. 4) при минимуме общего количества ошибок дискриминации.

Заключение

В данной работе показана возможность создания дискриминационных прогностических моделей для определения нефтяных разливов на суше с использованием снимков космического аппарата Landsat-8. Ожидается, что применение

методов атмосферной коррекции и использование исторических данных об анализируемых пикселях позволят получить модели с более высокой прогностической способностью.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Сайт Министерства природных ресурсов и экологии Российской Федерации [Электронный ресурс]: статьи. – Электрон. дан. – Режим доступа: <http://www.mnr.gov.ru/> – Загл. с экрана. – Яз. рус.
2. Википедия [Электронный ресурс]: статьи. – Электрон. дан. – Режим доступа: <http://ru.wikipedia.org/> – Загл. с экрана. – Яз. рус.
3. *Шовенгердт П.А.* Дистанционное зондирование. Модели и методы и обработки изображений: Учеб. пособие: Пер. с англ. А.В. Кирюшина, А.И. Демьяникова. – М.: Техносфера, 2010. – 556 с.
4. *Wold S., Esbensen K., Geladi P.* Principal Component Analysis // *Chemometrics and Intelligent Laboratory Systems*. – 1987, Vol. 2, P. 37-52.
5. *Sjöström M., Wold S., Lindberg W., Persson J.-Å., Martens H.* A multivariate calibration problem in analytical chemistry solved by partial least-squares models in latent variables // *Analytica Chimica Acta*. – 1983, Vol. 150, P. 61-70.
6. Сайт Геологической службы США [Электронный ресурс]: статьи, граф. дан. – Электрон. дан. – Режим доступа: <https://www.usgs.gov/> – Загл. с экрана. – Яз. англ.
7. Сайт The European Space Agency [Электронный ресурс]: статьи, граф. дан. – Электрон. дан. – Режим доступа: <http://www.esa.int> – Загл. с экрана. – Яз. англ.
8. Сайт Роскосмос [Электронный ресурс]: статьи. – Электрон. дан. – Режим доступа: <http://www.roscosmos.ru/> – Загл. с экрана. – Яз. рус.
9. *Bogomolov A.Yu., Galyanin V.V., Belikov V.A.* TPT cloud: a full-featured chemometrics software online // Winter Symposium on Chemometrics. Modern Methods of Data Analysis (10-я международная научная конференция), 2016. – <http://wsc.chemometrics.ru/media/files/conferences/wsc10/documents/WSC-10%20abstract%20book.pdf>.
10. *Galyanin V.V., Belikov V.A., Melentieva A.A., Bogomolov A.Yu.* Simulation of spectral data for regression methods validation // Winter Symposium on Chemometrics. Modern Methods of Data Analysis (10-я международная научная конференция), 2016.

Статья поступила в редакцию 2 февраля 2017 г.

ANALYSIS OF EARTH REMOTE SENSING DATA FOR OIL SPILL DETECTION

V.A. Belikov, V.V. Galyanin, S.P. Orlov

Samara State Technical University
244, Molodogvardeyskaya st., Samara, 443100, Russian Federation

The paper deals with the problem of oil spills detection caused by man's production activity. The use of a chemometrics web-service for build PCA, PLS-DA models for identify oil components on the surface is proposed. Examples of application of the models for petroleum products detection on the earth's surface in the Samara region are considered. It is shown that the separation of oil products, water resources, city buildings, vegetation and other soil formations obtained is highly accurate. With the help of cross-validation, a quantitative estimate of the guilt of the models used was made.

Keywords: modeling, web-service, Earth remote sensing data, cluster analysis, regression analysis.

*Vadim A. Belikov, Postgraduate Student.
Vladislav V. Galyanin, Head of Laboratory of Multivariate Analysis and Global Modelling.
Sergey P. Orlov (Dr. Sci. (Techn.)), Professor.*