

Применение методов автоматизированного дешифрирования в задачах распознавания зон антропогенного воздействия на нефтегазовых месторождениях

Элешкевич А.Д., Еременко М.С., Сайбель Е.Г., Христюлов И.А., Чернов А.Г.

АО «ТомскНИПнефть», Томск, Россия

nipineft@tomsknipi.ru

Аннотация

В данной статье рассматривается опыт создания и применения методов автоматизированного дешифрирования с использованием алгоритма машинного обучения для решения задач выявления зон антропогенного воздействия на территориях нефтегазовых месторождений. Описываются варианты использования различных исходных данных дистанционного зондирования Земли для выявления таких зон на примере лесных вырубок, анализируются их преимущества и недостатки. Предлагается оригинальный подход на основе нейронных сетей для дешифрирования аэрокосмических снимков и рассматриваются перспективы его использования.

Материалы и методы

Метод численного моделирования с использованием данных дистанционного зондирования Земли.

Ключевые слова

распознавание образов, данные ДЗЗ, антропогенные зоны, компьютерное зрение, машинное обучение, лесные вырубки, экологический мониторинг, нейронные сети

Для цитирования

Элешкевич А.Д., Еременко М.С., Сайбель Е.Г., Христюлов И.А., Чернов А.Г. Применение методов автоматизированного дешифрирования в задачах распознавания зон антропогенного воздействия на нефтегазовых месторождениях // Экспозиция Нефть Газ. 2023. № 7. С. 127–131. DOI: 10.24412/2076-6785-2023-7-127-131

Поступила в редакцию: 14.11.2023

AUTOMATION

UDC 528.77 | Original Paper

Application of automated decryption methods in recognition of anthropogenic impact zones on oil and gas fields

Eleshkevich A.D., Eremenko M.S., Saibel E.G., Khristolubov I.A., Chernov A.G.

“TomskNIPneft” JSC, Tomsk, Russia

nipineft@tomsknipi.ru

Abstract

This article discusses the experience of creating and implementing methods of automated decoding using machine learning algorithms to solve the problem of identifying areas of anthropogenic impact in oil and gas fields. The options for using various remote sensing data are described to identify such areas, using examples of forest clearings, along with their advantages and disadvantages. An original approach based on neural networks for decoding aerospace imagery is proposed, and its prospects for use are considered.

Materials and methods

Numerical modeling method using Earth remote sensing data.

Keywords

pattern recognition, remote sensing data, anthropogenic zones, computer vision, machine learning, forest cuttings, environmental monitoring, neural networks

For citation

Eleshkevich A.D., Eremenko M.S., Saibel E.G., Khristolubov I.A., Chernov A.G. Application of automated decryption methods in recognition of anthropogenic impact zones on oil and gas fields // Exposition Oil Gas, 2023, issue 7, P. 127–131. (In Russ).

DOI: 10.24412/2076-6785-2023-7-127-131

Received: 14.11.2023

Введение

Нефтяная индустрия, как одна из ведущих отраслей мировой экономики, обеспечивает глобальное энергетическое потребление и является ключевым фактором в формировании современного облика мировой инфраструктуры. Однако вместе с этими значительными выгодами важно оценивать возможные экологические и социальные проблемы, связанные с антропогенным воздействием на окружающую среду. Антропогенное воздействие в нефтяной сфере подразумевает загрязнение окружающей среды выбросами парниковых газов, нефтяными разливами, а также изменение природного ландшафта. Снизить воздействие возможно при его обнаружении на ранней стадии. В частности, современные технологии, процессы цифровизации и пространственный анализ (например, анализ аэрокосмических снимков) могут играть важную роль в распознавании и классификации антропогенных зон.

По мере расширения своей деятельности нефтегазовые компании часто вынуждены вырубать лесные массивы для создания необходимой инфраструктуры, в том числе для бурения скважин и строительства коммуникаций.

Одним из инструментов для контроля и наблюдения за состоянием окружающей среды является использование данных дистанционного зондирования Земли (ДЗЗ), которые при использовании геоинформационных систем наряду с методами автоматизированного дешифрирования данных ДЗЗ позволяют реализовать эффективный инструмент для контроля и борьбы с антропогенными угрозами экологии.

Однако распознавание образов по данным ДЗЗ представляет собой сложную и ресурсоемкую задачу. Необходимо учитывать различные факторы, такие как сезонные особенности (весенние поводки, временные обмеления), параметры съемки (процент облачности, число спектров в изображении, пространственное разрешение). Важным аспектом такой задачи является точность и скорость обработки больших объемов данных.

Актуальной задачей является то, что в условиях санкций коммерческие продукты (Erdas imagine, ArcGIS и другие) по автоматизированному дешифрированию стали недоступны, а это значит, что есть необходимость в создании импортонезависимой системы автоматизированного дешифрирования объектов на территориях нефтегазовых месторождений. Она позволит снизить трудозатраты сотрудников, так как в данный момент определение антропогенных зон происходит вручную. Кроме того, возрастает скорость обработки данных.

В АО «ТомскНИПнефть» сотрудниками управления экологии и специализированного института по геоинформационным системам совместно была поставлена задача по автоматизации процесса распознавания лесных вырубок как одного из факторов антропогенного воздействия на основании аэрокосмических снимков.

В рамках статьи приводится начальный этап работы по созданию системы автоматизированного дешифрирования зон антропогенного воздействия с примером в виде вырубок, а именно рассмотрены различные варианты данных ДЗЗ, а также построен один из алгоритмов для решения поставленной задачи.

О возможности применения различных данных ДЗЗ

В последние годы подход по распознаванию образов по результатам съемки беспилотных летательных аппаратов (БПЛА) активно развивается и масштабируется. Несмотря на это, исторически используемый подход по дешифрированию по космическим снимкам сохраняет ряд преимуществ, в том числе позволяет охватить большие территории и получить общий вид состояния лесных массивов на глобальном уровне. По космическим снимкам можно систематически наблюдать изменения в массиве лесов, а также анализировать динамику и воздействие на экосистему.

Определение области вырубок по космическим снимкам представляет собой сложную задачу, которая требует тщательного анализа при обработке данных. Основные сложности состоят в том, что вырубки, как правило, сопровождаются негативными факторами, такими как укрытие растительностью или недостаточное пространственное разрешение снимков. Космические снимки обладают ограничением на допустимое пространственное разрешение, что ярко выражено при работе с удаленными участками. Это может затруднять точную идентификацию вырубки и оценку соответствующих параметров, таких как размеры и форма.

Кроме того, сезонные изменения в растительности, уровень осадков или процент облачности влияют на качество и доступность космических снимков. Для получения достоверных результатов необходимо проводить анализ нескольких снимков, сделанных в разные временные периоды, например, в разные сезоны года. Это позволит учесть негативные факторы и получить более точные оценки.

БПЛА, как сравнительно новая технология, также предлагает ряд преимуществ для определения вырубок. Полученные снимки обладают высоким разрешением и позволяют определить не только границы вырубки, но и состояние, степень повреждения и тип вырубки. Это может быть полезно для небольших участков или труднодоступных районов, которые затруднительно определить с помощью космических снимков.

Кроме того, использование БПЛА обеспечивает возможность проведения детального анализа и дополнительных исследований, таких как сбор дополнительных данных о типах деревьев, структуре древостоя и здоровье леса. Это позволяет получать подробный вид вырубки и сделать анализ последствий для окружающей среды.

Однако, несмотря на преимущества БПЛА, использование данных аэрофотосъемки для распознавания области вырубок пока имеет существенный недостаток — высокую стоимости съемки. Для использования снимков БПЛА необходим ряд условий: средства БПЛА, подготовленные сотрудники, необходимое оборудование для разных видов съемки, командирование на места съемки. В итоге стоимость получается в разы выше, чем покупка космических снимков, что делает их ограниченными в использовании для некоторых организаций и исследователей. Кроме того, есть ограничения законодательства и требования в области безопасности, что снижает возможность их использования на некоторых территориях.

Использование методов машинного обучения

Определение антропогенных зон относится к классу задач по распознаванию образов. Для решения таких задач используются математические методы, например, методы пороговой обработки изображений, а также методы наращивания областей, методы машинного обучения [1, 2]. Последнее получило применение в области распознавания изображений благодаря способности обрабатывать большие объемы данных и находить сложные закономерности. Применение нейронных сетей для распознавания антропогенных зон может повысить точность и скорость выполнения этого процесса.

Преимуществом такого подхода является способность работать с неструктурированными данными, такими как космические снимки. Нейронные сети могут использовать пиксельные данные изображений и находить в них закономерности. Это способствует нахождению скрытых признаков или шаблонов, которые могут быть невидимы при анализе другими методами.

Одной из возможных структур нейронных сетей для распознавания являются сверточные нейронные сети (CNN — Convolutional Neural Networks) [3], изначально разработанные для обработки изображений и применяемые в анализе и классификации объектов. Сверточные слои нейронной сети извлекают базовые признаки, такие как линии, углы и текстуры, а потом распознают сложные уровни признаков, например, формы деревьев или границы вырубок. Это делает их подходящим выбором для обработки данных ДЗЗ.

Другой тип нейронных сетей, который может использоваться, — это рекуррентные нейронные сети (RNN — Recurrent Neural Networks) [4]. Они подходят для анализа последовательностей данных и могут использоваться для определения динамики вырубок. Например, могут предсказывать, как изменяются антропогенные зоны со временем или выявлять повторяющиеся закономерности в рассматриваемой области.

Использование нейронных сетей и машинного обучения в распознавании вырубок по данным ДЗЗ дает ряд преимуществ. Во-первых, автоматизирует процесс и повышает скорость обработки больших объемов данных, что позволяет сократить трудозатраты сотрудников на ручное распознавание, тем самым высвобождая время для других задач. Во-вторых, обнаруживает сложные закономерности и шаблоны, которые могут быть пропущены человеком, что может позволить достичь более высоких показателей точности, а также извлечь новую информацию для дальнейшего исследования. Наконец, нейронные сети и машинное обучение предполагают создание моделей, которые могут адаптироваться и улучшаться по мере получения новых данных, что положительно скажется на преимуществе такого подхода.

Описание построенного алгоритма дешифрирования

В рамках статьи рассматривается использование космических снимков для распознавания вырубок в силу большей практической применимости. Такое решение обуславливается вышеописанными преимуществами и недостатками различных данных ДЗЗ.

Ключевым моментом для распознавания вырубок является качество исходных данных (параметры разрешения, зашумленность изображения, точность разметки

и масок), на основе которых будет происходить обучение и валидация модели. Для исследовательских целей в обучении модели обычно используются данные из открытых источников, однако для решения производственных задач было принято решение использовать более качественные ранее закупленные и предварительно распознанные оператором космические снимки. Спутниковые снимки имеют разрешение до 2-х метров с 4-мя каналами и 16-битной цветовой гаммой. Данные космические снимки были предварительно подготовлены для обучения модели: размечены с учетом сезонности, растительности и других факторов. Примеры исходных данных и их маски можно увидеть на рисунках 1 и 2.

Стоит отметить, что показанные вырубки в виде транспортной инфраструктуры и площадных объектов составляют основную часть имеющегося набора данных. В обучающей выборке, необходимой для обучения нейронной сети, использовались космические снимки общей площадью 22,89 км², из которой вырубки составляют 2,94 км².

Задача распознавания вырубок по космическим снимкам в нашей постановке относится к классу задач семантической сегментации, когда каждому пикселю входного изображения должен быть подобран свой класс. В данном случае речь о бинарной классификации.

В рамках исследуемой задачи нами было принято решение использовать базовую структуру нейронной сети с многослойным перцептроном. Решение поставленной задачи было начато с рассмотрения его возможностей. Несмотря на то, что перцептрон не способен учесть сложные взаимосвязи между данными, такие как геометрическое расположение пикселей относительно друг друга, у него есть некоторые преимущества, обусловленные простой структурой и небольшим числом параметров:

- перцептрон обладает высокой скоростью обучения;
- перцептрон обладает возможностью интерпретации влияния параметров на конечный результат;
- перцептрон требует меньше вычислительных ресурсов для обучения и сохранения модели;

• перцептрон справляется с задачами, где данные содержат простую структуру.

Для передачи данных для обучения в перцептрон применялся подход на основе конвертации многомерного массива данных космического снимка в двухмерный массив, чтобы значения каналов одного пикселя содержались в одной строке, а сами номера строк представляли собой номера исходных пикселей (рис. 3).

Для обучения представленных данных была построена полносвязная нейронная сеть с двумя скрытыми слоями 32 и 64 нейрона в каждом. Входным слоем служило число нейронов, соответствующее числу каналов входного изображения, в нашем случае 4. Выходной слой представлял собой два нейрона для бинарной классификации (рис. 4.). Функция активации для скрытых слоев была выбрана Relu, а для выходного слоя Softmax [5]. В качестве метода оптимизации был выбран алгоритм Adam [6], в качестве функции потерь была принята кросс-энтропийная разница между маской и прогнозами модели, а в качестве метрики при обучении использовалась точность. Обучение прекращалось, если точность модели при прохождении эпохи не увеличилась больше, чем 0,001.



Рис. 1. Площадные объекты и их маски, как вырубки
Fig. 1. Area objects and their masks, like cuttings



Рис. 2. Транспортная инфраструктура и ее маска
Fig. 2. Transport infrastructure and its mask

Текущие результаты

Модель показала существенную зависимость от входных данных для обучения и валидации. Ввиду самостоятельного выбора порога значения вероятности для бинарной классификации можно говорить только об интервальном оценивании характеристик модели. При применении обученной нейронной сети на данных, полученных с тех же космических аппаратов и того же сезона съемки, что и в обучающей выборке, результаты по точности и полноте находятся в диапазоне 71–78 %. В то время как для существенно отличающихся по характеристикам разрешения или сезону съемки космических снимков результаты распознавания являются низкими и находятся в диапазоне 18–26 % точности и полноты. Эту проблему можно решить путем увеличения выборки данных или использования более сложных нейронных сетей.

На рисунке 6 представлен результат работы нейронной сети для предложенных

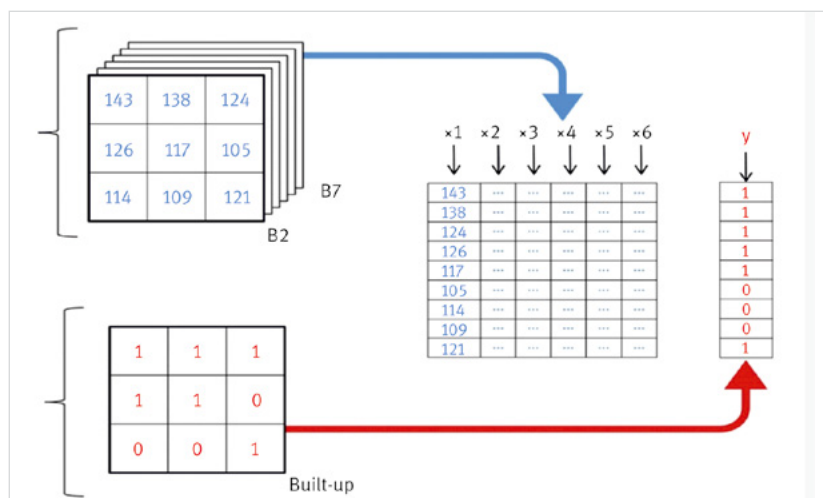


Рис. 3. Принцип конвертации массивов данных
Fig. 3. Principle of converting data sets

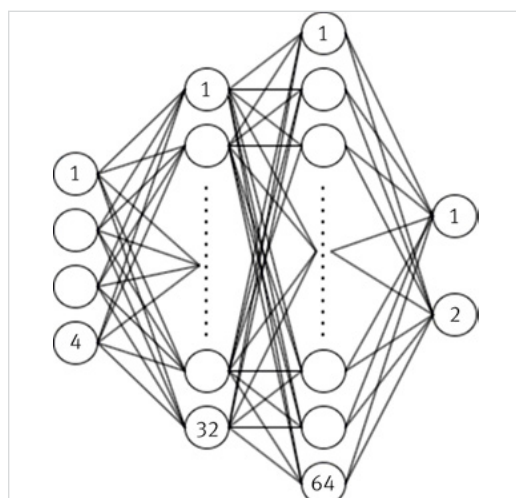


Рис. 4. Схема нейронной сети
Fig. 4. Neural network diagram

космоснимков (рис. 5). Чем ближе цвет к черному, тем выше вероятность, что это вырубка.

Итоги

В дальнейшем планируется применить другие алгоритмы автоматического дешифрирования, например, нейронные сети типа U-net [7], DeepLabV3 [8], ResNet [9] и другие, а также предусмотреть метрики для сравнения результатов работы алгоритмов. Дополнительно необходимо рассмотреть интеграцию алгоритмов распознавания образов с импортозамещающей платформой. Результатом всех этих исследований будет являться система по автоматизированному дешифрированию зон антропогенного воздействия для решения производственных задач АО «ТомскНИПИнефть» для решения производственных задач, на базе которой в дальнейшем можно будет рассматривать не только зоны антропогенного воздействия, но и другие задачи геопространственного анализа.

Выводы

На основании проведенного исследования можно сказать, что задача автоматического

дешифрирования зон антропогенного воздействия актуальна в рамках экологического мониторинга на нефтегазовых месторождениях. Для её выполнения можно использовать различные виды данных ДЗЗ, такие как космические снимки и снимки БПЛА. Оба варианта имеют преимущества и недостатки, однако в силу экономических обоснований и большего практического использования на данный момент космические снимки представляют больший интерес. Однако при снижении стоимости применения БПЛА и облегчения законодательного регулирования их использование в будущем может дать более точные результаты. Использование подхода на основе нейронных сетей к задаче определения антропогенных зон может сократить трудозатраты сотрудников и предоставить возможность сфокусироваться на анализе уже полученных результатов.

Литература

1. Кривогуз Д.О. Анализ изменения землепользования и растительного покрова (LULC) Керченского полуострова с 1990 по 2020 гг. с помощью ДЗЗ и машинного обучения //

Экспериментальное ландшафтоведение: теория, методология, практика. 2023. С. 36–37.

2. Напряшкин А.А. Алгоритмическое и программное обеспечение системы интерпретации аэрокосмических изображений для решения задач картирования ландшафтных объектов. Диссертация. Томск: 2002. 183 с.
3. Багаев И.И. Анализ понятий нейронная сеть и сверточная нейронная сеть, обучение сверточной нейросети при помощи модуля TensorFlow // Математическое и программное обеспечение систем в промышленной и социальной сферах. 2020. Т. 8. № 1. С. 15–22.
4. Кострюкова М.И. Распознавание цветных изображений с использованием рекуррентных нейронных сетей в сравнении со сверточными нейронными сетями // Актуальные проблемы прикладной математики, информатики и механики. Воронеж: Вэлборн, 2022. С. 537–543.
5. Соснин А.С., Сулова И.А. Функции активации нейросети: сигмоида,



Рис. 5. Участок космоснимка для распознавания
Fig. 5. Satellite image area for recognition



Рис. 6. Результат работы алгоритма
Fig. 6. Result of algorithm

- линейная, ступенчатая, relu, tahn // Наука. Информатизация. Технологии. Образование. Екатеринбург: РГППУ, 2019. С. 237–246.
6. Петухов Д.Е., Белов Ю.С. Обзор часто используемых алгоритмов по оптимизации стохастического градиентного спуска // E-Scio. 2021. № 1. С. 553–561.
7. Alexakis E., Armenakis C. Evaluation of UNet and UNet++ architectures in high resolution image change detection applications. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, Vol. 43, P. 1507–1514. (In Eng).
8. Yang W., Zhang J.L., Xu Z.Y., Hu K. A real-time DeepLabv3+ for pedestrian segmentation. Opticheskiy Zhurnal, 2019, Vol. 86, issue 9, P. 49–59. (In Eng).
9. Рудь А.А., Рудь С.А. Сравнение архитектур сверточных нейронных сетей на основе Resnet в задаче распознавания изображений // Проблемы техники и технологий телекоммуникаций ПТИТТ-2020. Самара: ПГУТИ, 2020. С. 305–306.

ENGLISH

Results

In the future, it is planned to explore other algorithms for automatic decryption, such as neural networks like Unet [7], DeepLabV3 [8], ResNet [9], and others, as well as provide metrics for comparing the performance of these algorithms. Additionally, integration of image recognition algorithms with import-substituting platforms should be considered. The outcome of these studies will be a system for automated decryption of anthropogenic impact zones, which will be implemented in “TomskNIPIneft JSC” to address production tasks. Moreover, this system can serve as a basis for further analysis of not only anthropogenic impact zones but also other geospatial analysis tasks.

References

1. Krivoguz D.O. Analysis of land use and land cover change (LULC) of the Kerch Peninsula from 1990 to 2020 using remote sensing and machine learning. Experimental landscape science: theory, methodology, practice. 2023. P. 36–37. (In Russ).
2. Napryushkin A.A. Algorithmic and software systems for interpreting aerospace images for solving problems of mapping landscape objects. Dissertation. Tomsk: 2002, 183 p. (In Russ).
3. Bagaev I.I. Concept analysis neural network and conventional neural network, conventional neural network training using the TensorFlow module. Software of systems in the industrial and social fields, 2020, Vol. 8, issue 1, P. 15–22. (In Russ).

Conclusions

Based on the conducted research, it can be stated that the task of automatic decryption of anthropogenic impact zones is relevant within the framework of environmental monitoring in oil and gas fields. Various types of remote sensing data can be used for this purpose, such as satellite imagery and unmanned aerial vehicle (UAV) images. Both options have advantages and disadvantages, however, due to economic justifications and greater practical use at the present time, satellite imagery is of greater interest. However, as the cost of UAV application decreases and legislative regulations become more lenient, their use in the future may produce more accurate results. Employing a neural network-based approach to the identification of anthropogenic zones can reduce the labor requirements of personnel and provide the opportunity to focus on the analysis of the obtained results.

4. Kostryukova M.I. Recognition of color images using recurrent neural networks in comparison with convolutional neural networks. Current problems of applied mathematics, computer science and mechanics. Voronezh: Welborn, 2022, P. 537–543. (In Russ).
5. Sosnin A.S., Suslova I.A. Functions of neural net activation: sigmoid, linear, step, relu, tahn. The science. Informatization. Technologies. Education. Ekaterinburg: Russian State Vocational Pedagogical University, 2019, P. 237–246. (In Russ).
6. Petukhov D.E., Belov Yu.S. Review of commonly used algorithms for optimizing stochastic gradient descent. E-Scio, 2021, issue 1, P. 553–561. (In Russ).
7. Alexakis E., Armenakis C. Evaluation of UNet and UNet++ architectures in high resolution image change detection applications. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, Vol. 43, P. 1507–1514. (In Eng).
8. Yang W., Zhang J.L., Xu Z.Y., Hu K. A real-time DeepLabv3+ for pedestrian segmentation. Opticheskiy Zhurnal, 2019, Vol. 86, issue 9, P. 49–59. (In Eng).
9. Rud A.A., Rud S.A. Comparison of convolutional neural network architectures based on Resnet in the problem of image recognition. Problems of engineering and technology of telecommunications PTTT-2020. Samara: Volga State University of Telecommunications and Informatics, 2020, P. 305–306. (In Russ).

ИНФОРМАЦИЯ ОБ АВТОРАХ | INFORMATION ABOUT THE AUTHORS

Элешкевич Алексей Дмитриевич, студент 2 курса магистратуры, Томский государственный университет, Томск, Россия; инженер, АО «ТомскНИПИнефть», Томск, Россия
Для контактов: elshkevichad@tomsknipi.ru

Еременко Марк Сергеевич, инженер, АО «ТомскНИПИнефть», Томск, Россия

Сайбель Екатерина Георгиевна, ведущий инженер, АО «ТомскНИПИнефть», Томск, Россия

Христолюбов Илья Александрович, начальник отдела геоинформационных технологий, АО «ТомскНИПИнефть», Томск, Россия

Чернов Артем Геннадьевич, к.т.н. главный специалист, АО «ТомскНИПИнефть», Томск, Россия

Eleshkevich Alexey Dmitrievich, 2nd year master's student, Tomsk State University, Tomsk, Russia; engineer, “TomskNIPIneft” JSC, Tomsk, Russia
Corresponding author: elshkevichad@tomsknipi.ru

Eremenko Mark Sergeevich, engineer, “TomskNIPIneft” JSC, Tomsk, Russia

Saibel Ekaterina Georgievna, leading engineer, “TomskNIPIneft” JSC, Tomsk, Russia

Hristolyubov Ilya Aleksandrovich, head of the geoinformation technology department, “TomskNIPIneft” JSC, Tomsk, Russia

Chernov Artem Gennadievich, ph.d. of engineering sciences chief specialist, “TomskNIPIneft” JSC, Tomsk, Russia