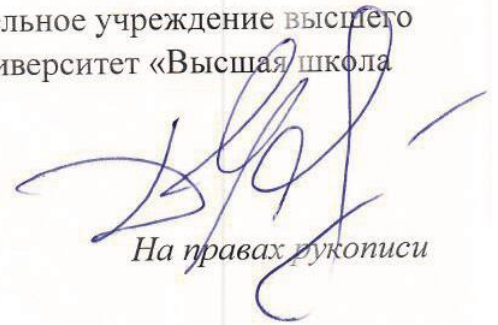


Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет «Высшая школа экономики»



На правах рукописи

Максименко Даниил Дмитриевич

**МЕТОДЫ ОПТИМАЛЬНОГО ВЫБОРА ЛОКАЦИИ ДЛЯ КОМПАНИЙ
ВЫСОКОТЕХНОЛОГИЧНОГО БИЗНЕСА НА ТЕРРИТОРИИ РЕГИОНОВ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

Специальность 5.2.3 – Региональная и отраслевая экономика
(региональная экономика)

Диссертация на соискание учёной степени
кандидата экономических наук

Научный руководитель:
доктор экономических наук, профессор
Фонотов Андрей Георгиевич

Москва – 2024

Оглавление

ВВЕДЕНИЕ.....	3
1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ РАЗМЕЩЕНИЯ ПРЕДПРИЯТИЙ.....	15
1.1 Эволюция в вопросах теории размещения.....	15
1.2 Современный этап развития теории размещения.....	23
1.3 Факторы размещения высокотехнологичных предприятий.....	36
1.4 Изменения в целеполагании бизнеса в России и за рубежом.....	45
2. РАЗМЕЩЕНИЕ ВЫСОКОТЕХНОЛОГИЧНЫХ ПРЕДПРИЯТИЙ НА ТЕРРИТОРИИ РЕГИОНОВ РОССИИ.....	53
2.1 Аспекты использования микроданных.....	53
2.2 Информационная база исследования.....	56
2.3 Анализ размещения ВТП России с применением микроданных.....	63
3. МЕТОДИЧЕСКИЕ АСПЕКТЫ РЕШЕНИЯ ЗАДАЧИ РАЗМЕЩЕНИЯ ВЫСОКОТЕХНОЛОГИЧНЫХ ПРЕДПРИЯТИЙ.....	74
3.1 Пути решения задач размещения.....	74
3.2 Интерактивные системы поддержки принятия решений.....	76
3.3 Использование машинного обучения для поиска благоприятной локации для ВТП.....	85
3.4 Формирование оптимальной модели выбора локации.....	99
ЗАКЛЮЧЕНИЕ.....	105
СПИСОК ЛИТЕРАТУРЫ.....	111
ПРИЛОЖЕНИЕ.....	127
Приложение 1. Блокнот с ходом машинного обучения для решения задачи размещения высокотехнологичного бизнеса.....	127

ВВЕДЕНИЕ

Актуальность темы диссертационного исследования¹. Размещение производства оказывает большое влияние на его дальнейшую деятельность на каждом этапе жизненного цикла. Выбор производственной площадки является одним из первых стратегических решений для предприятия, он определяет параметры доступности рынков сбыта, трудовых и производственных ресурсов, немалую часть операционных расходов, существенно влияет на себестоимость продукции. В случае с высокотехнологичными предприятиями (ВТП) вопрос размещения стоит еще более остро, так как для таких предприятий важен доступ к ряду специфических ресурсов, например, к рынку высококвалифицированных кадров, научной инфраструктуре и т.д.

Одним из последствий санкционного давления на Россию со стороны ключевых игроков на рынке высокотехнологичной продукции (США, Германия, Япония), которое привело к разрыву торгово-экономических связей с рядом зарубежных стран, нарушению цепочек поставок, закрытию ряда производств, стало существенное повышение актуальности задач развития отечественной промышленности. В настоящее время локальные перебои зарубежных поставок преодолеваются путем расширения товарной номенклатуры импорта из дружественных стран или параллельного импорта.

¹ Диссертация содержит фрагменты исследований, опубликованных автором в виде научных статей, включая, но не ограничиваясь: Максименко Д. Д. Актуальные факторы размещения высокотехнологичных производств в России // Региональные исследования. 2023. № 4. С. 29–40; Михайлов А. С. и др. Научно-инновационная динамика развития Северо-Запада России в условиях геополитических перемен // Балтийский регион. 2023. Т. 15, №. 4. С. 79–103; Mikhaylova A. et al. Geoinformation representation of maritime knowledge flows: new frontiers of coastalization // IOP Conference Series: Earth and Environmental Science. IOP Publishing. 2022. Vol. 1087, No. 1. P. 012038; Maximenko D., Maximenko M. GIS for location planning of banks' physical networks // Regional Studies, Regional Science. 2021. Vol. 8, No. 1. P. 362–365; НИР «Разработка интерактивной платформы по оценке состояния ключевых отраслей экономики с применением альтернативных источников статистики»; квалификационных работ “Location factors of high technological business in Moscow” и «Факторы размещения малых инновационных предприятий в Российской Федерации».

Однако для обеспечения технологического суверенитета в среднесрочной перспективе и наращивания объема отечественных высокотехнологичных товаров и услуг в полтора раза к 2030 году в соответствии с "Перечнем поручений по реализации Послания Президента Федеральному Собранию" (утв. Президентом РФ 30.03.2024) необходимо появление новых и развитие существующих центров высокотехнологичной промышленности, что является серьезным вызовом как для федеральных, так и для региональных властей.

Среди прочего для решения этих задач требуется оценка потенциала территории и формирование алгоритмов принятия решений для размещения производств на основе анализа факторов размещения, релевантных, с одной стороны, специфике сектора высоких технологий, а с другой – связанных с особенностями социально-экономического и историко-географического развития регионов России, а также современными технологическими и стратегическими вызовами.

На фоне стремительной цифровизации российской экономики существенно расширяется перечень информационных источников (как государственных, так и частных) для анализа актуальной территориальной структуры и потенциальных факторов размещения высокотехнологичных производств. Развитие государственных информационных систем в рамках Национальной программы «Цифровая экономика Российской Федерации», а также реализация государственной программы «Национальная система пространственных данных» позволяют проектировать цифровые инструменты для поддержки принятия решений при выборе оптимальной локации для ВТП на всей территории страны.

Степень научной разработанности темы диссертации. Одним из ключевых элементов при выборе локации предприятия являются факторы размещения – наиболее важные контекстные условия для развития бизнеса на определенной территории. Анализ факторов размещения проводится в научных и практических работах со времени Второй промышленной революции (конец XIX – начало XX вв.), когда размораживание производительных сил и ресурсов стало особенно серьезно влиять на успех предприятий.

Сам набор факторов эволюционировал с развитием производственных и рыночных отношений, появлением новых отраслей промышленности, усложнением организационных структур предприятий и их обособлением, повышением требований к условиям труда со стороны квалифицированного персонала и т.д.

Как итог первоначальные математические задачи оптимального размещения производств по отношению к рынку и ресурсной базе конца XIX в. на современном этапе сменились сложными мультимасштабными исследованиями, изучающими самый широкий набор факторов (от «классических» факторов наличия трудовых ресурсов и доступа к рынку до культурных и институциональных) как на уровне стран и макрорегионов, так и конкретных локаций внутри населенных пунктов.

Одной из первых и наиболее известных работ по изучению факторов размещения является «Изолированное государство» Иоганна фон Тюнена [147] (первоиздание от 1826 г.). Несмотря на свою простоту, концепция фон Тюнена стала прорывной для своего времени и сформировала почву для последующих исследований (в том числе и современных). Работа наглядно показывает, что размещение предприятия может отражаться на его эффективности (рента от размещения).

Дальнейшее развитие теория размещения получила в конце XIX – начале XX в. в Германии, где в продолжение идей фон Тюнена на фоне развития тяжелой промышленности и транспортной инфраструктуры ученые начали применять математические методы для обоснования размещения производств с целью снижения издержек [149].

В 1930-х гг. комплексные работы Вальтера Кристаллера и Августа Леша формируют теорию центральных мест и теорию экономики размещения [51, 111]. Эти направления ставили своей целью формирование оптимальных иерархических структур размещения предприятий на основе параметров рынков сбыта и транспортной сети.

В 1950–1960-е гг. изучение теории размещения активно развивается в англосаксонской науке. Наиболее значительными становятся исследования американских ученых, преследовавших прикладные цели размещения при изучении данного направления [83].

Параллельно развивался позитивистский подход в анализе размещения. Уолтер Айзард – один из наиболее известных ученых, проводивших количественные исследования размещения в 1950-х. В том или ином виде, базируясь на концепциях Вебера, в работе «Размещение и пространственная экономика» [85] он представляет субституционный подход к размещению предприятий.

Советская география в свою очередь также развивала собственные концепции выбора локации, исходя из особенностей плановой экономики: задачи размещения решались в рамках советской районной школы. В своих работах [16,23,24] советские ученые использовали деление страны на районы – цельные территориальные общности, связанные природными и социально-экономическими факторами.

В 1970–1990-х гг. теория размещения нашла развитие в обширном перечне научных школ и направлений, адаптируя подходы поведенческих наук, урбанистики и т.д. [91]. Наиболее влиятельным направлением этого периода является новая экономическая география (НЭГ) [72,73,99].

В современных работах по размещению предприятий [33,37,50] рассматривается широкий список различных показателей, используя методы из географии (региональный и картографический анализ), экономики (эконометрические методы), социологии (опросы).

Размещения высокотехнологичных и инновационных предприятий – одна из наиболее распространенных тем исследований теории размещения [89,92,93,143]. Высокотехнологичный бизнес формирует специфические требования к ресурсной базе территории – в первую очередь для ВТП важны качественные характеристики факторов размещения (трудовых ресурсов, социальной сферы).

Несмотря на достаточно высокую степень изученности темы анализа территориальных особенностей развития высокотехнологичного бизнеса [14,112] в России и за рубежом, недостаточно проработанным остается вопрос анализа факторов размещения высокотехнологичных производств. Кроме этого, в российском научном пространстве вопросы размещения анализируются на уровне регионов, что представляется недостаточным на фоне их географических масштабов и существования внутрирегиональных особенностей территорий, информационная база исследований зачастую ограничена данными официальной статистики.

Объектом исследования являются территориальные системы размещения высокотехнологичных предприятий. Под высокотехнологичными предприятиями в работе подразумеваются компании с основными ОКВЭД2 21 (фармацевтика) и 26 (электроника), 20 (химия), 27 (электрооборудование), 28 (машины), 29 (автотранспорт) [76,110]*.

Диссертационное исследование выполнено в соответствии со следующими пунктами Паспорта научной специальности 5.2.3. «Региональная и отраслевая экономика (региональная экономика)»:

1.2. «Пространственная организация национальной экономики. Пространственное распределение экономических ресурсов» в части оценки актуальной территориальной структуры и факторов размещения высокотехнологичных предприятий;

1.16. «Оценка и прогнозирование перспектив развития региональных экономических систем» в части формирования методики для оценки потенциала территорий с целью размещения новых высокотехнологичных предприятий.

Предметом исследования являются факторы, влияющие на размещение высокотехнологичных предприятий, и методы их учета при выборе локации.

* То есть в периметр анализа наряду с высокотехнологичными включены среднетехнологичные предприятия высокого уровня на уровне второго знака ОКВЭД2

Среди ключевых гипотез исследования – превалирование влияния факторов научно-образовательной среды в перечне основных факторов размещения ВТП и важная роль особых экономических зон при выборе локации.

Цель и задачи исследования. Цель диссертационного исследования – разработка системы методов и инструментов для оценки благоприятности локации для высокотехнологичного предприятия на территории регионов России.

В ключевые задачи исследования входят:

1) анализ научных источников по теме теории размещения с целью формирования перечня факторов размещения, релевантных для высокотехнологичного бизнеса и методов выбора локации;

2) подготовка информационно-аналитической базы для выбора локации с использованием микро- и пространственных данных, а также инструментов бизнес-аналитики;

3) изучение актуальной конфигурации размещения высокотехнологичного бизнеса и формирование типологии стратегических решений при выборе локации;

4) формирование алгоритма выбора локации для высокотехнологичного бизнеса, в том числе с использованием инструментов машинного обучения.

Методология исследования. Методологически исследование делится на несколько последовательных этапов. На первом был проведен библиографический анализ литературы, посвященной тематике размещения бизнеса, факторам размещения и современным инструментам анализа размещения предприятий.

На основе библиографического анализа сформирован массив факторов, актуальных для анализа размещения ВТП, рассмотрена возможность количественного анализа факторов на основе доступных в России наборов данных. Кроме того, проведен историко-экономический и историко-географический анализ драйверов размещения ВТП в России и за рубежом.

В ходе исследования изучены инструменты и методы, применяемые при анализе размещения и выборе локации академическим и бизнес-сообществом. Была подготовлена географическая карта конфигурации высокотехнологичного

комплекса России с использованием микроданных финансовой отчетности высокотехнологичных предприятий. Анализ ареалов размещения предприятий, а также примеров функционирования отдельных площадок позволил сформировать типологию бизнеса по основным мотивам выбора локации в формате архетипов.

Работа с большими массивами микроданных, а также пространственной информации позволила сформировать пошаговый алгоритм создания интерактивной платформы поддержки принятия решений как одного из возможных вариантов оптимизации экспертной работы.

На основе предшествующих этапов сформирована модель машинного обучения, оценивающая потенциал территории для размещения высокотехнологичных предприятий на основе квантифицированных факторов размещения на уровне муниципалитетов и крупных агломераций. Выявлены факторы размещения, имеющие наибольший вес при моделировании.

По результатам исследования и апробации различных методов выбора локации подготовлена концептуальная модель действий при оценке потенциала территории для размещения ВТП.

Научная новизна. Научная новизна диссертационного исследования заключается в разработке системного подхода к анализу факторов размещения высокотехнологичных предприятий и его адаптации в рамках создания инструментов поддержки принятия управленческих решений при планировании привлечения ВТП на территорию региональными и федеральными органами власти. Новизна раскрывается в следующих положениях, выносимых на защиту:

1. Выявлены факторы размещения, наиболее применимые для анализа территориальной структуры высокотехнологичных предприятий;
2. Выявлены ключевые ареалы размещения ВТП на основе визуализации пространственных данных. Выделены архетипичные стратегии размещения предприятий российского высокотехнологичного сектора;
3. Разработан и апробирован подход к интеграции микроданных из разных источников на базе платформы бизнес-аналитики для решения задач размещения высокотехнологичных предприятий в России;

4. Разработана и апробирована модель машинного обучения для решения задач размещения высокотехнологичных предприятий на уровне муниципалитетов и крупных агломераций;
5. Сформирован алгоритм принятия решений о размещении предприятий, использующий последовательное применение апробированных методологических инструментов. Сформировано несколько модульных подходов к решению вопросов размещения.

Теоретическая значимость диссертационного исследования.

Теоретическая значимость диссертационного исследования заключается в разработке междисциплинарного системного подхода к анализу факторов и методов выбора локации высокотехнологичных предприятий.

В современной российской экономической и географической повестке вопрос факторов размещения рассмотрен недостаточно подробно, особенно в приложении к высокотехнологичному комплексу; работы по теме в основном базируются на данных регионального масштаба, из-за чего малоизученной остается внутри- и межрегиональная специфика.

Анализ ареалов распространения высокотехнологичных предприятий позволяет выделить отдельные факторы размещения, актуальные для ВТП России на текущий момент. Во многом выявленные факторы соотносятся или дополняют выводы, описанные в классических и современных работах [17, 100, 153] по размещению производств (центр – периферийный разрыв, наличие «эффекта колеи», значимость специальных экономико-производственных зон). Однако также выявлен малоизученный фактор обслуживающей функции высокотехнологичных предприятий по отношению к менее технологичным отраслям локальной специализации (нефтегазовому сектору, металлургии, энергетике и т.д.), который формирует новые центры развития ВТП, снижая территориальные диспропорции развития высокотехнологичного сектора [19].

Практическая значимость исследования. Практическая значимость диссертационного исследования заключается в разработке алгоритмов принятия решений и оценке значимости отдельных факторов размещения ВТП. Результаты

исследования могут быть востребованы как со стороны государственной власти для оценки потенциала и поиска решений для развития высокотехнологичных предприятий на территории, так и со стороны бизнеса для формирования набора локаций, привлекательных для размещения предприятий.

Таким образом, результаты диссертационного исследования в первую очередь актуальны для ФОИВов и РОИВов с позиций анализа пространственных сдвигов размещения ВТП с целью подготовки, имплементации и мониторинга стимулирующих решений для расширения присутствия ВТП на территории. Также разработки в рамках диссертации могут быть полезны крупному бизнесу (в том числе компании с государственным участием): апробированные в работе сочетания методических подходов позволяют выстраивать стратегию размещения региональной сети компании, основанную на данных, что потенциально снижает риск принятия ошибочных управленческих решений.

Объективность и достоверность. Среди прочего работа основывается на объективных микроданных финансовой отчетности предприятий, базах открытых данных с доступным набором метаданных. По ходу работы прописаны вероятные ограничения в использовании информационных источников, периметре и методологии исследования. Исследовательские выводы и методологические подходы, сформированные по ходу подготовки диссертации, валидированы в рамках процесса рецензирования статей автора, написанных во время исследовательской работы в рамках аспирантуры.

Апробация результатов диссертационного исследования. В первую очередь апробация промежуточных материалов исследования осуществлялась в рамках подготовки научных статей:

- 1) Максименко Д. Д. Актуальные факторы размещения высокотехнологичных производств в России // Региональные исследования. 2023. № 4. С. 29–40, (ВАК К1) [19]. В статье проводится анализ факторов размещения, актуальных для высокотехнологичных предприятий в России, а также рассматривается территориальная структура ВТП, выявляются типичные стратегии размещения подобных предприятий;

- 2) Maximenko D., Maximenko M. GIS for location planning of banks' physical networks // *Regional Studies, Regional Science*. 2021. Vol. 8, No. 1. P. 362–365, (Scopus Q2 в рубрике “Economics and Econometrics”) [118]. В статье рассматривается применимость использования инструментов пространственного анализа и ГИС для научно-практических задач размещения, вклад автора – 50%;
- 3) Михайлов А. С., Максименко Д. Д. и др. Научно-инновационная динамика развития Северо-Запада России в условиях геополитических перемен // *Балтийский регион*. 2023. Т. 15, № 4. С. 79–103, (ВАК К1) [21]. В работе анализируются глобальные изменения в научно-технологической повестке на фоне изменения структуры стран - партнеров России, апробируется использование данных ЕГИСУ НИОКТР для мониторинга научно-практического ландшафта регионов, вклад автора – 25%;
- 4) Михайлов А. С., Максименко Д. Д., Максименко М. Р. Пространственные и структурные закономерности в распределении научно-технологической, инновационной и производственной деятельности в России // *Балтийский регион*. 2024. Т. 16. № 2. С. 41–62, (ВАК К1) [22]. В работе исследуется система связей инновационного, научного и производственного потенциала регионов России, вклад автора – 33%.
- 5) Mikhaylov A., Mikhaylova A., Maximenko D., Maximenko M., Hvalej D. Coastal regions in the geography of innovation activity: a comparative assessment of marine basins // *Geographica Pannonica*. 2022. Vol. 26, No. 4. P. 345–355, (Scopus Q2 в рубрике “Geography, Planning and Development”) [125]. В работе анализируется влияние прибрежного и приграничного положения на научно-исследовательский потенциал территории, вклад автора – 25%;
- 6) Mikhaylova A., Maximenko D. et al. Geoinformation representation of maritime knowledge flows: new frontiers of coastalization // *IOP Conference Series: Earth and Environmental Science*. IOP Publishing. 2022. Vol. 1087, No. 1. P. 012038, [126]. В статье рассматривается территориальная структура системы

трансфера знаний с применением визуализации в формате карты потоков – 25%.

Наработки, подготовленные в ходе исследования, также были представлены на российских и международных конференциях:

- 1) II Международный статистический форум СНГ: новые технологии и источники данных в статистике, 2024 год, Ташкент «Разработка интерактивной платформы по оценке состояния ключевых отраслей экономики с применением альтернативных источников статистики» (докладчик);
- 2) XXIII Апрельская конференция, 2022 год, Москва. Д. Д. Максименко, М. Р. Максименко «Разработка базы микроданных государственных закупок лекарственных средств для анализа фармацевтических рынков в России» (докладчик);
- 3) XXIV Апрельская конференция, 2023 год, Москва. Д. Д. Максименко, М. Р. Максименко «Использование микроданных для анализа науки и технологий: пример России» (докладчик);
- 4) IWEG 2022, 2022 год, Циндао. Mikhaylova A., Maksimenko D., Hvalej D., Mikhailov A., Maksimenko M. Geoinformation representation of maritime knowledge flows: new frontiers of coastalization.

Помимо научных статей и выступлений на конференциях, в ходе подготовки диссертационного исследования: сбора данных и апробации методологического аппарата, были подготовлены четыре внутренних результата интеллектуальной деятельности (РИД) НИУ ВШЭ:

- 1) РИД № 8.0231-2022 «Методика выгрузки и агрегирования пространственных данных (в том числе на основе адресов организаций) с применением ГИС-инструментов на примере закупок лекарственных средств», вклад автора – 25%;
- 2) РИД № 8.0127-2023 «Методология сбора и предобработки первичных микроданных для анализа отраслей реального сектора экономики», вклад автора – 25%;

- 3) РИД № 8.0126-2023 «Разработка архитектуры баз данных для специализированных отраслевых аналитических платформ», вклад автора – 25%;
- 4) РИД № 8.0128-2023 «Формирование интерактивной платформы отраслевой аналитики на облачной платформе», вклад автора – 25%.

Сформированные в диссертации подходы по созданию системы поддержки принятия решений для размещения бизнеса были использованы при подготовке научно-исследовательской работы (НИР) «Разработка интерактивной платформы по оценке состояния ключевых отраслей экономики с применением альтернативных источников статистики», вклад автора – 25%.

Структура и объем диссертации. Диссертационное исследование состоит из введения, трех глав, включающих одиннадцать разделов, заключения, списка литературы и приложения с результатами моделирования. Основные разделы посвящены:

- 1) литературному анализу, а также обзору российского и международного опыта выбора локации для высокотехнологичных предприятий, оценке значимости отдельных групп факторов размещения по отношению к ВТП;
- 2) использованию микро- и пространственных данных для формирования систем поддержки принятия решений о размещении бизнеса;
- 3) картированию ВТП на основе данных корпоративной отчетности и анализу архетипичных стратегий размещения предприятий в России;
- 4) апробации использования машинного обучения для решения задач оценки потенциала территории для выбора локации и формированию концептуальной модели действий при выборе локации.

Общий объем работы – 126 страниц без учета приложения.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ РАЗМЕЩЕНИЯ ПРЕДПРИЯТИЙ

1.1 Эволюция в вопросах теории размещения

Среди основополагающих теоретических подходов работы – подходы теории размещения, междисциплинарного направления науки, изучающего вопросы выбора локации предприятиями и влияние этого выбора на их деятельность. За более чем 200 лет своего существования теория размещения сформировала обширный методологический аппарат, опирающийся на концепции географии, экономики, социологии. Современный инструментарий (использование ГИС, открытых данных) позволяет изучать вопросы размещения на разных масштабных уровнях с высокой точностью.

Одной из первых и наиболее известных работ по изучению факторов размещения является «Изолированное государство» Иоганна фон Тюнена [147]. В своей модели представления реальности он показал гипотетическое идеальное государство с аграрной экономикой, состоящее из города и пригородной части. В зависимости от удаления от центра этого государства разные экономические активности сменяют друг друга так, что каждая расположена наиболее эффективно в зависимости от транспортных издержек, потребительских свойств продукции и ее цены – таким образом формируется дополнительная рента от размещения. Например, производство молока должно быть расположено ближе к центральному городу, так как продукт должен быть быстро потреблен: как следствие возможности по транспортировке ограничены (Рисунок 1).

Несмотря на свою простоту, концепция фон Тюнена стала прорывной для своего времени и сформировала почву для последующих исследований (в том числе и современных). Работа наглядно показывает, что верное размещение предприятия может отражаться на его эффективности (рента от размещения). Кроме того, работа «предсказывает» существование агломерационного эффекта, в

связи с чем часто упоминается последователями новой экономической географии [65].

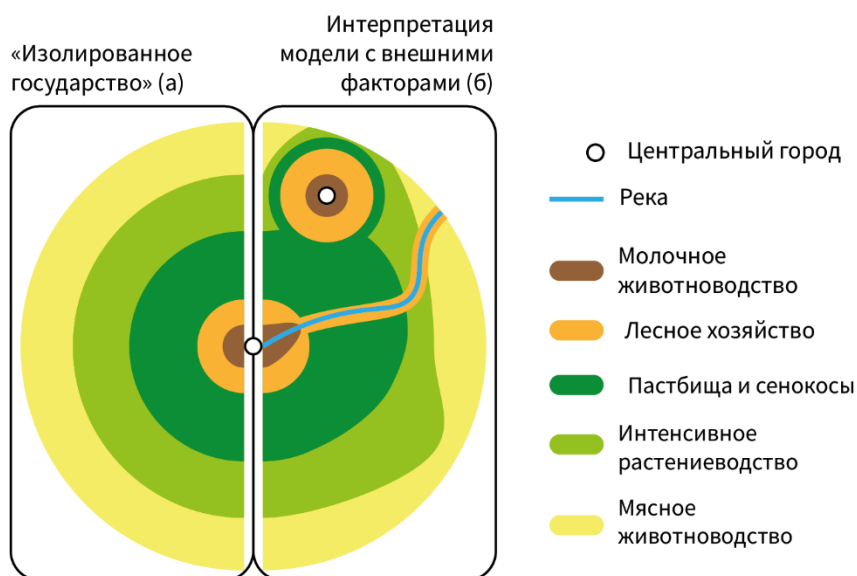


Рисунок 1 – Модель фон Тюнена (а) и ее возможная интерпретация (б)

Источник: The Geography of Transport Systems [136].

Дальнейшее развитие теория размещения получила в конце XIX – начале XX в. в Германии, где в продолжение идей фон Тюнена на фоне развития тяжелой промышленности и транспортной инфраструктуры ученые начали применять математические методы для обоснования размещения производств с целью снижения издержек.

Подобные работы обычно использовали подходы геометрии для описания экономических связей между размещением сырьевой базы, производства и потребителя. Например, в «Теории размещения промышленности» [149] Альфред Вебер использует «штандортную фигуру» для оптимизации размещения предприятия между рынком сбыта и двумя источниками сырья (Рисунок 2). В 1930-х гг. комплексные работы Вальтера Кристаллера и Августа Леша формируют теорию центральных мест [51] и теорию экономики размещения [111]. Эти направления ставили своей целью формирование оптимальных иерархических структур размещения предприятий на основе доступности рынков сбыта и транспортных артерий.

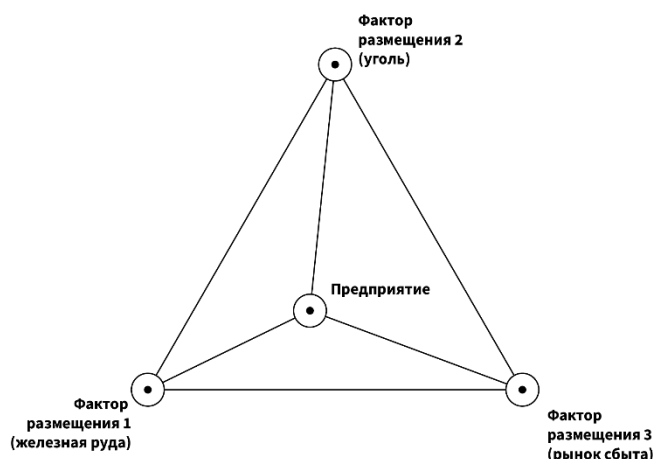


Рисунок 2 – «Штандортная фигура» А. Вебера

Источник: Leontief Technology and the Location of the Firm in a Weber Triangle—Specific Localization Theorems [101].

Немецкие исследования сформировали огромный теоретический базис для развития теории размещения: использование математических методов для обоснования размещения стало важным методом не только для науки, но и для практических целей. Однако геометрические модели имели ряд ограничений, основным из которых является серьезное усложнение при добавлении дополнительных факторов для анализа.

В неоклассической экономической теории важность размещения изучал Альфред Маршалл: в частности, в работе «Принципы экономики» [117] 1890 г. он описал концентрацию производств как один из факторов снижения издержек. Среди позитивных сторон концентрации он отмечает [13]:

1) повышение профессионального уровня персонала за счет «перетоков» навыков и знаний (несмотря на то, что эта концепция не упоминается, она отражает идею автора);

2) более быстрое распространение новых технологий и процессов за счет их обсуждения внутри профессионального сообщества – «диффузия инноваций»;

3) появление вспомогательных отраслей;

4) более высокий уровень загрузки производственных мощностей при однородности локализованного производства;

5) появление стабильного рынка труда для квалифицированных работников.

А. Маршалл также отметил недостатки процессов локализации:

1) повышение издержек на размещение в центральных районах, поскольку они становятся более привлекательными из-за своих преимуществ (доступ к рынку и кадрам);

2) рынок труда имеет высокий уровень однородности (высокая конкуренция за трудовые ресурсы), однако это негативное влияние может быть нивелировано с развитием вспомогательных отраслей;

3) если локализация представлена фирмами с одинаковой специализацией, снижение спроса на их товары или проблемы с поставками сырья могут нанести ущерб экономическому процветанию региона. Тем не менее формирование нескольких ведущих отраслей может ослабить потенциальное кризисное состояние.

Работа Маршалла важна для теории размещения: она интегрирует это направление с экономической теорией, а некоторые идеи, заложенные в «Принципах...» (эффекты концентрации) актуальны и для современных работ: как и в случае с фон Тюненом, работа Маршалла предсказала многие популярные сейчас в экономике концепции.

В 1950–1960-х гг. лидерство в изучении теории размещения переходит к англосаксонской науке. Наиболее успешными становятся исследования американских ученых, преследовавших практические цели в изучении данного направления. Идеи в работах американских авторов представляют широкий набор анализируемых факторов размещения [43], основанных не только на транспортно-логистических функциях, но и на качестве городского пространства, человеческого капитала, политическом участии и т.д.

Пример такого подхода – работа профессора Гарвардского университета Эдварда Гувера «Размещения экономических активностей». Среди основных факторов размещения в работе отмечены городское планирование, землепользование и т.д. Важный вывод работы – помимо экономических выгод, существуют также социальные экстерналии от размещения. Гувер утверждает,

что факторы размещения не статичны, постоянно меняются, так как и территория и сами предприятия меняют свои параметры (миграция населения, уровень квалификации рабочих, размер фирмы) со временем, поэтому оценка размещения должна быть максимально гибкой [83].

Параллельно на фоне совершенствования математических методов развивался позитивистский подход в анализе размещения. Уолтер Айзард – один из наиболее известных ученых, проводивших количественные исследования размещения в это время. В том или ином виде, базируясь на концепциях Вебера, в работе «Размещение и экономия от пространства»: [85] он представляет субституционный подход к размещению предприятий. Исходя из этого подхода, разные факторы могут «подменять» друг друга (например, большая цена труда может быть нивелирована меньшими транспортными издержками), поэтому теоретическое оптимальное размещение находится на кривой линии по орбите от рынка потребления, пересекающей источники сырья.

Советская география в свою очередь также развивала собственные концепции размещения. Для СССР было критически важной задачей оптимизировать размещение новых промышленных предприятий на фоне форсированной индустриализации, поэтому задачи размещения производств носили практический характер – найти оптимальные позиции для промышленных комплексов на огромной территории страны.

Задачи размещения решались в рамках советской районной школы, среди ключевых представителей этого направления – Николай Николаевич Баранский, Николай Николаевич Колоссовский, Юлиан Глебович Саушкин и Анатолий Тимофеевич Хрущев. В своих работах [16,23,24] они использовали деление страны на районы – цельные территориальные общности, связанные природными и социально-экономическими факторами. Каждый район обладает своим набором ресурсов (в том числе человеческим капиталом и его качеством), на основе которого можно понять размещение каких производств в нем наиболее рационально. Кроме того, формировались территориально-промышленные комплексы (ТПК) – наиболее благоприятные наборы отраслей на основе их

технологической и экологической комплексности, подходящие для территории на основе ее особенностей и ресурсов.

Наиболее интересный аспект советской теории размещения – кооперация науки и реальных проектировочных решений. Советские географы анализировали потенциальные размещения как качественные (анализируя экспедиционные материалы), так и количественные (рассчитывая существующие и перспективные экономические балансы) методами. Важно заметить, что несмотря на предубеждения человеческий капитал играл одну из лидирующих ролей среди других факторов размещения в советских исследованиях. Например, в таких работах анализировались потенциальные издержки на создание качественной инфраструктуры для привлечения квалифицированных кадров.

Однако достижения советской теории размещения применимы по большей части только для плановой экономики: предприятия в рамках ТПК создают сложную систему связей, каждая единица в которой (даже убыточная индивидуально) важна для всего комплекса – поддержание таких систем в рыночной экономике является достаточно сложной задачей.

Сравнивая западные и советские исследования той эпохи, важно отметить, что несмотря на усложнение подходов теории размещения ключевые факторы новых моделей оставались фундаментальными – транспортные издержки, свойства и размещение сырья, потребители, хотя к ним и добавился человеческий капитал, так как новые индустрии потребовали привлечения наиболее квалифицированных кадров. Концепция ТПК стала продвинутой практической и теоретической платформой для решения задач пространственного и экономического развития СССР. Напротив, западные исследователи были более гибкими, анализируя широкий спектр факторов на разных масштабных уровнях, они в том числе пытались выявить уже существующие факторы принятия решений о размещении бизнесом.

Позднесоветские и российские ученые вели работу по адаптации идей районирования и промышленных комплексов под новые реалии: среди прочего применительно к российской действительности были интерпретированы идеи М.

Портера, П. Кругмана, И. Валлерстайна. В их статьях предпринимаются попытки повышения роли социально-экономических, управленческих и даже культурологических аспектов в рассмотрении районов и различных понятийных надстроек над концепцией ТПК [99, 131, 148]. Однако на фоне экономических, а главное – политических изменений и конъюнктуры научного большинства для управления территорией и, как следствие, оценки факторов потенциала территории для размещения бизнеса широкое распространение получили идеи регионалистики, региональной политики и экономики. Так, Александр Григорьевич Гранберг причисляет вопросы теории размещения к области региональных исследований и региональной экономики в частности [9].

В 1970–1990-х гг. вопросы теории размещения нашли развитие в обширном перечне научных школ и направлений, адаптируя подходы поведенческих наук, урбанистики и т.д. [56, 91, 141]. Так, в рамках бихевиористского подхода выбора локации поднимается вопрос фактора иррациональности такого выбора в связи с ограничениями, связанными с неполнотой данных и/или отсутствием возможности их адекватной интерпретации, предубеждениями и «ментальными картами» (искажениями в восприятии территории), а также фактором неопределенности [119].

Наиболее важным для исследования направлением этого этапа является новая экономическая география (НЭГ), сформированная американским экономистом Полом Кругманом в конце 1980-х гг. Эта школа представляет факторы размещения через комплексную систему, разделяя их на факторы первой (физическая география, климат) и второй (человеческий капитал) природы [72,73,99]. Ключевой идеей этой экономической школы (несмотря на «географию» в названии) является существование возрастающих доходов от концентрации ресурсов и инновационных процессов, формирующих центр-периферийные отношения между территориями разного ранга. НЭГ смогла эмпирическими методами доказать существование пространственных закономерностей, выявленных в классических работах о теории размещения.

Некоторыми учеными [48] выделяется направление анализа размещения фирм в рамках подходов бизнес-демографии. Согласно такому подходу жизненный цикл компании делится на ряд условных этапов: в основном это «рождение» (например, из фазы стартапа), слияние и поглощение, переезд и «смерть». Изучение компаний в совокупности в пространственно-временном разрезе позволяет рассматривать благоприятность территории (социально-экономического ландшафта) для размещения или развития бизнеса, или, наоборот, выявлять факторы и события, предшествующие закрытию фирм. Одним из наиболее действенных методов в рамках подхода является анализ hot-spot, т.е. анализ тепловых карт (или их наложения) в приложении к концентрации создания или ликвидации фирм.

Проводя обобщающий анализ исторического развития теории размещения, можно сказать, что это направление науки решало современные для каждого отдельного исторического этапа проблемы экономики. Базовая модель фон Тюнена была приемлема для аграрного общества с серьезными лимитирующими факторами в виде транспортных издержек и ограничений по объемам хранения продукции. Индустриализация и появление железных дорог поставили более сложные вопросы перед теорией размещения: периметр возможного размещения заметно расширился вместе с повышением его важности, ведь речь шла в том числе и о крупных промышленных предприятиях.

Теория размещения доказала свою эффективность на практике и стала одним из инструментов регионального промышленного развития как в России, так и за рубежом. Усложнение экономических связей заставляло применять новые методы и рассматривать новые факторы размещения. НЭГ привнесла новые идеи в концепции в теорию и эмпирически доказала существование неявных закономерностей размещения, углубив идеи кластеризации, концентрации и диффузии инновации, развивая теорию размещения как актуальное направление для постиндустриального общества.

1.2 Современный этап развития теории размещения

В современных работах по анализу локации бизнеса рассматривается широчайший список различных показателей [31], используя методы из географии (региональный и картографический анализ), экономики (эконометрические методы), социологии (опросы) и т.д. Вместе с тем их методологическая и концептуальная разношерстность не позволяет объединять их по каким-либо направлениям или научным школам. Лучшие практики из теории размещения используются в геомаркетинге и региональной политике.

Среди ключевых стимулов для локации ВТП в современных исследованиях отмечается образовательная инфраструктура [109,114]. Связь между локацией и географической близостью к университетам с помощью метода наименьших квадратов (МНК) просчитана в работе «Перетоки знаний университетов и размещение новых фирм» [37]. Цель этой работы – выяснить роль перетоков знаний со стороны университетов в принятии фирмами решений о размещении. Авторы представляют обширный обзор литературы, в котором подтверждается идея о том, что перетоки знаний от университетов пространственно ограничены, а их сила коррелирует с научным потенциалом вузов и исследовательских институтов. В статье описываются два механизма осуществления перетоков знаний:

- 1) перетоки знаний через научное знание. В данном случае трансфер идей происходит через кодифицированное знание в формате публикаций и научных исследований, т.е. потенциал трансфера знаний можно измерить через публикационную активность. Предполагается, что с развитием интернет-библиотек, репозиториев и онлайн-коммуникаций значимость географической близости в отношении этого механизма перетока знаний снизилась, причем больше для естественно-научных направлений, чем для социальных наук;
- 2) перетоки знаний через человеческий капитал. По мере того, как выпускники учебных заведений переходят в бизнес-среду, они передают

некодифицированные знания, которые могут быть ценными для конкурентоспособности высокотехнологичных предприятий. Для этого механизма перетоков пространственная близость может быть очень важна, поскольку неявная передача знаний требует живого общения. В этом случае потенциал трансфера знаний может быть измерен по количеству учащихся.

Тем не менее сила эффекта перетока знаний со стороны университетов может варьироваться из-за особенностей конкретной локации, например, мобильности студентов, стоимости аренды вблизи университетов, отраслевой принадлежности фирм-реципиентов знаний и т.д. Основываясь на данных о высокотехнологичных фирмах и университетах Германии, авторы подготовили количественный анализ факторов размещения. Результаты этого анализа свидетельствуют о том, что перетоки знаний пространственно ограничены, их влияние наиболее отчетливо выявляется в пределах 7 км от вузов. Тем не менее и более отдаленные от университетов фирмы, получают отдельные конкурентные преимущества в доступе к новым знаниям и высококвалифицированным работникам.

Одним из интересных направлений в работе с теорией размещения является география инноваций [137]. Концепции географии инноваций популярны в современных работах по анализу размещения высокотехнологичных отраслей, так как они раскрывают дополнительные мотивы и нестандартные факторы их размещения.

Например, в работе Мартина Кенни и Дональда Пэттона «География предпринимательства: сети поддержки трех высокотехнологичных индустрий» [93] авторы анализируют стратегии размещения стартапов в сфере биотехнологий, полупроводников и телекоммуникаций в США на базе их концентрации и близости к «агентам поддержки», таким как:

- 1) юридические фирмы. Они отмечены как один из важнейших участников жизненного цикла стартапов, поскольку имеют дело с интеллектуальной собственностью, которая аккумулирует наибольшую ценность венчурных фирм;

2) венчурные инвесторы. Так как бизнес стартапов является рискованным, венчурные капиталисты пытаются сохранить свои инвестиции, принимая участие в развитии фирм. Например, они могут знакомить потенциальных потребителей с продукцией стартапов, представлять руководителей стратегическим партнерам и т.д. Опытные венчурные капиталисты предпочитают работать с местными стартапами, чтобы более эффективно контролировать их деятельность;

3) инвестиционные банкиры. Роль банкиров становится более важной на зрелой стадии стартапа, когда он готов к IPO, поэтому их роль в размещении неравномерна. В США инвестиционные банки в значительной степени сосредоточены в Нью-Йорке, за некоторыми исключениями в Силиконовой долине и Бостоне;

4) частные инвесторы. Эта группа агентов обладает широким спектром намерений и знаний: например, они могут быть возможными потребителями товаров стартапов или оказывать консультационные услуги в развитии проекта, будучи компетентными в необходимой отрасли.

Чтобы выяснить пространственную важность поддерживающей инфраструктуры для каждой отрасли, пары фирм и их поддерживающих агентов были разделены по критерию близости (более или менее 50 миль). Проведенный анализ хи-квадрат (для проверки гипотезы о пространственной концентрации) показал, существуют ли статистические различия в близости к поддерживающим агентам между отраслями или нет.

Работа доказывает существование неоднородности в мотивах принятия решений о размещении между различными высокотехнологичными отраслями. Исследование также является примером сочетания качественного с простым, но эффективным количественным анализом, основанным на пространственных факторах.

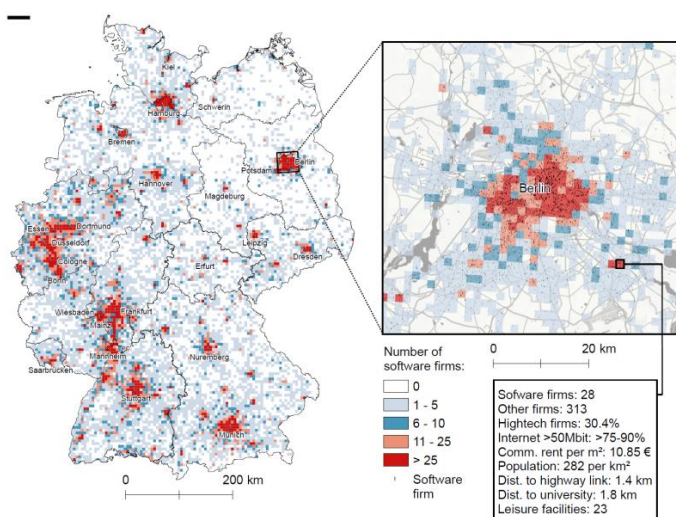


Рисунок 3 – Пример визуализации пространственных данных по размещению фирм

Источник: Analyzing and predicting micro-location patterns of software firms [96].

Одна из наиболее актуальных и релевантных для темы исследования работ – «Анализ и предсказание закономерностей размещения производителей программного обеспечения на микроуровне» [96]. Авторы используют комбинацию эконометрических и картографических методов для анализа размещения (Рисунок 3). Как сама идея, так и качество реализации показывают реальность анализа факторов размещения фирм на разных масштабных уровнях, используя концепции теории размещения в связке с современными инструментами ГИС и эконометрики.

Работа носит в основном методологический характер, поскольку в ней не ставится цель описать факторы размещения или объяснить мотивы принимаемых решений, а оценивается применимость возможных методов пространственного и эконометрического анализа для решения задач оценки потенциала территории для выбора локации.

Во-первых, в статье исследуются результаты агрегирования географических данных на разных уровнях масштаба для регрессионного моделирования: это актуальный вопрос в современных пространственных исследованиях – проблема модифицируемых площадных единиц (MAUP – Modifiable Area Unit Problem).

МАУР устанавливает, что объединение точечных объектов в площадные для анализа может существенно отразиться на результаты последующего анализа как из-за масштаба сформированных площадных объектов, так и из-за их формы. Во-вторых, в статье представлены методы прогнозирования размещения фирм-разработчиков программного обеспечения, основанные на анализе пространственных данных с применением эконометрического моделирования.

Основной особенностью этой работы является то, что она использует с пространственные и статистические данные на микроуровне. Авторы утверждают, что анализ на более высоких уровнях агрегации (регионы, округа и т.д.) может значительно исказить интерпретацию влияния отдельных факторов размещения. Подобные исследования стали доступны благодаря наличию микроданных: как географической информации, собираемой на волонтерских началах картографическим сообществом, так и официальных (открытых) геоданных.

Результаты регрессий, подготовленных в рамках работы, показали влияние следующих факторов размещения:

1) агломерационные факторы. В эту группу факторов входят плотность населения и концентрация бизнеса, стоимость коммерческой аренды. Каждый из факторов показал важность для размещения фирм-разработчиков программного обеспечения с положительными коэффициентами. Квадраты плотностей, добавленные в модель, показали u-образное влияние концентрации, а также высокий уровень значимости в модели, но с отрицательным коэффициентом: это означает, что фирмы-разработчики ПО избегают ядер агломераций;

2) инфраструктурные факторы. Меньшие расстояния до ближайших автомагистралей, железных дорог и аэропортов, а также наличие доступа к широкополосному подключению к Интернету важны для работы модели. Доступ к общественному транспорту не показал значимости;

3) социально-экономические факторы размещения. К ним относятся близость к ближайшему университету, более низкий уровень налогообложения и

более высокая доля квалифицированной рабочей силы. Они также оказались важными факторами размещения IT-фирм;

4) факторы размещения, влияющие на качество жизни. Ожидаемая продолжительность жизни и количество объектов для отдыха оказались важными предикторами размещения IT-фирм, в то время как криминогенная обстановка, количество объектов культуры и досуга, не имели значимости для модели;

5) прочие факторы. Использование цифровой модели рельефа местности и переменной с оценкой качества геокодирования значительно повысили качество модели.

Проверка качества работы модели на разных масштабных уровнях выявила аномалии, неточности в оценках концентрации ВТП внутри агломераций. Использование модели для прогнозирования на более высоком уровне агрегации снижает эффект пространственной автокорреляции и приводит к меньшему количеству ошибок.

Описанная выше работа является примером разностороннего исследования, демонстрирующего квалифицированный подход к подготовке базы данных, выбору методов анализа, визуализации и интерпретации результатов. Поскольку в работе анализируются IT-компании, некоторые использованные в ней факторы могут быть применены и для высокотехнологичных фирм России. Статья также указывает на важность качества и количества данных для анализа вопросов размещения: не все наборы использованных данных имеют аналоги для территории России на представленных масштабных уровнях.

Продолжает тему анализа наукоемких предприятий, но уже на уровне города работа «Соседство инноваций в городе: закономерности размещения инновационных фирм на микроуровне в Берлине» [133]. В этой статье исследуются пространственные механизмы диффузии инновации внутри агломерации. Авторы рассматривают влияние контекстных объектов на размещение инновационного бизнеса на базе пространственных микроданных.

По этой причине авторы используют неагрегированные пространственные данные, чтобы избежать проблему MAUP. В работе применяются концепции

руководства Осло по определению инновационных фирм и конкретных видов инновационной активности.

Интересной особенностью этой работы является то, что в ней используются данные «Берлинской инновационной панели» за 5-летний период, чтобы выявить, влияет ли бизнес-демография инновационных фирм на инновационный климат в регионе. Авторы предполагают, что размещение инновационных предприятий представляет собой процесс с возможной эндогенной природой: фирмы стремятся расположиться в районе с подходящими условиями, но и их размещение само по себе изменяет окружающую среду.

Результаты сравнительного анализа выявили разницу между инновационными и неинновационными фирмами. Для инноваторов характерно меньшее расстояние до государственных научно-исследовательских институтов и университетов, они, как правило, расположены ближе к железнодорожной сети и в менее населенных районах. Тем не менее расстояние до объектов досуга и показатель цен на недвижимость не влияют на выбор локации. Hot-spot анализ для определения различий между инновационными и неинновационными фирмами показал наличие специализированных инновационных кластеров (например, рядом с Берлинским техническим университетом).

Работа показывает закономерности размещения инновационных фирм на уровне города, выявляя существование эффектов перетока знаний, которые могут стимулировать инновационную активность в рамках определенного географического охвата.

Еще одним распространенным типом работ по тематике являются антологии факторов размещения [43, 116]. За неимением единой школы ученые собирают результаты исследований своих коллег в сборники, которые могут послужить базой для гипотез и дизайна последующего исследования. Пример такой работы – «Измерения факторов размещения: обзор и объяснение» [33]. Ее цель – систематизировать существующие факторы размещения и оценить их ценность для современных индустрий.

К исследовательским вопросам работы относятся следующие: какие факторы размещения существуют, на какие группы их можно разделить и какие из них применимы для принятия решений о размещении международных компаний. Автор использует библиографический обзор по основным работам в области теории размещения и формирует на их основе список из 205 факторов размещения, разделенный на 14 тематических групп (включая климатические, экономические, политические, социальные факторы и т.д.). После этого список был сокращен до 100 факторов в результате экспертной оценки и валидации выпускниками университетов США и ОАЭ со специализацией в области экономики, операционного менеджмента и географии.

На следующем этапе анализа был проведен опрос 2125 промышленных компаний различного размера (количество крупных, средних и малых предприятий в выборке примерно одинаково) из 25 стран. Результаты опроса показали, что еще 5 факторов размещения из 100 не имеют отношения к мотивам размещения современными отраслями промышленности (это близость к религиозным объектам и библиотекам, доступность почтовых услуг и т.д.).

Эта работа раскрывает неоднородность представлений о факторах размещения среди ученых и практиков. Тем не менее результаты работы показывают, что даже неочевидные факторы размещения принимаются во внимание международными компаниями, поскольку они имеют решающее значение для их успеха, особенно для новых филиалов в зарубежных странах.

Описанный набор работ показывает разнообразие методов, масштабов анализа и исследовательских вопросов, стоящих перед современной теорией размещения. Есть и другие работы, рассмотренные в ходе изучения научной литературы, которые не были включены в подробный анализ, их основные характеристики показаны в Таблице 1.

Таблица 1 – Обзор современных работ по теме размещения предприятий

Работа	Территория (объект) анализа	Техники	Важные факторы размещения
Business micro-location: factors, preferences and indicators to assess the influence factors [78]	Порто (Португалия) и Бухарест (Венгрия)	Использование комбинации матриц со взвешенными результатами межотраслевого опроса о мотивах выбора локации	9 ключевых факторов, сформированных на основе обзора теоретических концепций: близость к рынку сбыта, арендная плата и т.д.
The odd role of proximity in knowledge relations: high-tech in the Netherlands [145]	Нидерланды (региональный уровень)	МНК регрессия на региональном уровне, зависимая переменная – количество высокотехнологичных фирм и сотрудников	Доля технических специалистов/бакалавров на региональном рынке труда, наличие агломерации в регионе, количество университетов, количество неакадемических институтов
Beyond the Silicon Valley: University R&D and high-technology location [153]	Соединенные Штаты Америки (уровень графств)	Логит-модель для сверхдисперсных данных, зависимая переменная – количество новых высокотехнологичных заводов	Затраты на рабочую силу, стоимость недвижимости, налоги, размер рынка, наличие квалифицированных кадров, затраты на НИОКР и т.д.

<p>Location Theories and Business Location Decision: A Micro-Spatial Investigation of a Nonmetropolitan Area in Canada [64]</p>	<p>Низовье Святого Лаврентия (административный район Квебека, Канада)</p>	<p>Модель дискретного выбора</p>	<p>Экономическое разнообразие, удаленность от агломераций, размер фирм, а также их специализация и территориальная концентрация</p>
<p>Science Parks and the growth of new technology-based Firms — academic-industry links, innovation and markets [109]</p>	<p>Швеция</p>	<p>Регрессионный анализ для выявления влияния научных парков на занятость и уровень доходов фирм</p>	<p>Научные (технологические) парки</p>
<p>How knowledge links with universities may foster innovation: The case of a science park [60]</p>	<p>Научные парки Мадрида</p>	<p>Анализ социальных сетей и траекторий для выявления неявных связей между деятельностью университетов и фирм внутри научных парков</p>	<p>Университеты могут стимулировать инновационную активность (создавать перетоки знаний) технологических и научных парков</p>
<p>Proximity, networking and knowledge production in Europe: What lessons for innovation policy? [116]</p>	<p>Европа (региональный уровень)</p>	<p>МНК и пространственная авторегрессия, зависящая от переменной — количество патентов на душу населения</p>	<p>Географическая, технологическая и социальная близость</p>

The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms [143]	Соединенные Штаты Америки	Пуассоновская и отрицательная биномиальная регрессии	Территориальная концентрация фирм, университетов, патентов и венчурного капитала
The macro- and micro-environmental factors of decisions of production facility location by Japanese companies in Poland [150]	Нижняя Силезия	Интервью в фокус-группе с проведением 5-балльной оценки факторов размещения	Налоги, стоимость рабочей силы и инвестиций, близость к рынкам
University research and the location of business R&D [28]	Великобритания (региональный уровень)	Регрессионный анализ на региональном уровне, зависимая переменная – количество бюро НИОКР	Колокация исследовательских бюро ВТП и вузов
Tradeoffs between local taxes and government spending as determinants of business location [75]	Штат Мэн, Соединенные Штаты Америки (уровень графств)	Пуассоновская и отрицательная биномиальная регрессии	Налоговая нагрузка как положительный фактор, отражающийся на состоянии социально-экономической среды
Competition for business location: A survey [59]	–	Обзор теоретических концепций и теоретическое моделирование	Региональная конкуренция посредством налоговых ставок и субсидирования бизнеса, технологическое развитие региона, размер рынка, инфраструктура

Evaluating the importance of business location factors: The influence of facility type [95]	Соединенные Штаты Америки	Опрос по 4-балльной шкале Лайкерта о важности списка из 39 факторов размещения и последующий дисперсионный анализ (ANOVA) и тест Тьюки	Близость к аэропорту, стоимость жилья, стоимость рабочей силы, качество школьного образования и т.д.
Firm-specific and location-specific drivers of business location and relocation decisions [41]	–	Обзор литературы	Существование push (внутренняя мотивация на переезд фирмы) и pull (внешние преимущества от размещения или переезда фирмы на территории) факторов релокации бизнеса
Location and organizing strategy: Exploring the influence of location on the organization of pharmaceutical research [74]	Соединенные Штаты Америки и Европейский Союз	Историко-географический и историко-экономический анализ, МНК регрессии	Географическая близость научно-технологической базы

Источник: составлено автором.

Таким образом, тематика анализа размещения остается привлекательной для исследователей, в том числе для таких лидеров науки, как Пол Кругман и Рон Бошма. Среди ключевых направлений современных исследований можно выделить:

- работы, изучающие факторы размещения предприятий [44,100,109]. В них оценивается влияние отдельных объектов (университетов, транспортной

инфраструктуры и т.д.) и явлений (инновационной активности предприятий на территории, институциональной среды и т.л.) на вероятность размещения бизнеса, в том числе с помощью опроса экспертного сообщества;

- работы по прогнозированию размещения предприятий. В них рассматриваются инструменты прогнозирования локации бизнеса на основе релевантных факторов размещения [96];
- обзоры литературы и метаанализы, в которых агрегируются факторы размещения, релевантные для конкретных типов предприятий (например, высокотехнологичных или транснациональных). Основой таких работ является библиографический анализ, который может быть дополнен экспертной валидацией [40].

В российской науке исследования в рамках теории размещения (в том числе в приложении к высокотехнологичным отраслям) проводятся сравнительно редко. Исключениями являются работы коллективов РАНХиГС [14], МГУ им. М.В. Ломоносова [17] и ИЭОПП СО РАН [27], ИЭ УрО РАН [18] серия работ по теме размещения и адаптации транснациональных корпорации в России подготовлена И.Б. Гурковым [10] в НИУ ВШЭ – разноплановость исследовательских коллективов указывает на междисциплинарную природу вопросов размещения. При этом, в научном поле обширны наработки в смежных тематических направлениях, напрямую влияющих на вопросы размещения ВТП – по умной специализации [12], географии инноваций [6], вопросам импортозамещения [1].

В то же время, в отличие от зарубежных статей, инструментальный аппарат, набор данных для анализа и масштабный ряд российских исследований ограничен. В первую очередь рассматриваются либо региональный, либо локальный уровень в формате анализа примеров отдельных предприятий – внутрирегиональные аспекты размещения системно в работах не представлены. Из этого следует и другая особенность – ограниченный инструментарий эконометрических оценок, связанный с априорно малым

числом наблюдений. Также достаточно редко встречается использование микроданных (в том числе опросных) в дезагрегированном виде – для консистентности данных их объединяют по сетке регионально административно-территориального деления [27].

1.3 Факторы размещения высокотехнологичных предприятий

В Таблице 2 приведены ключевые факторы и группы факторов размещения, представленные в современных научных источниках, их специфика по отношению к ВТП и источник данных, через который эти факторы можно квантифицировать для количественной оценки степени их влияния на размещение высокотехнологичных предприятий. Далее дана оценка их актуальности для размещения ВТП в России.

Таблица 2 – Факторы размещения предприятий

Группа факторов	Факторы и контекстные объекты	Актуальность для высокотехнологичных производств	Актуальность для России	Источники	Квантификация
<i>Научная инфраструктура</i>	Университеты	Возникновение перетоков знаний, рекрутинг квалифицированных кадров, трансфер технологий	Высокая	[28, 31, 145]	Данные мониторинга высшего образования
	Исследовательские институты		Высокая	[17, 74]	Данные ЕГИСУ НИОКТР
	Бюро НИОКР		Высокая	[28, 105, 153]	
<i>Человеческий капитал и качество жизни</i>	Уровень заработной платы	Генерация и привлечение квалифицированных трудовых ресурсов	Высокая	[96]	Данные Росстата
	Уровень преступности		Средняя	[90, 94, 138]	Портал правовой информации
	Стоимость жизни		Средняя		Данные Росстата
	Рекреационная инфраструктура		Средняя		Данные Open Street Map
	Медицинская инфраструктура		Средняя		Данные Росздравнадзора
	Образовательная среда		Средняя		Данные мониторинга высшего образования/данные

					рейтингового агентства РАЕХ/данные Министерства культуры России
<i>Производственная и логистическая инфраструктура</i>	Технопарки, гринфилды и браунфилды	Удобство развертывания производства, сокращение издержек и кооперация поставок	Высокая	[57, 60, 109]	Данные Минпромторга России
	Транспортная инфраструктура		Высокая	[4, 78, 121]	Данные Open Street Map
	Инженерная инфраструктура		Высокая		Данные Open Street Map
	Стоимость недвижимости		Средняя		Данные отраслевых агрегаторов (ЦИАН)
	Стоимость воды, электроэнергии и др.		Средняя		
<i>Средовые факторы</i>	Социально-экономическое состояние территории	Привлекательность территории для инвестиций и ведения бизнеса	Высокая	[71, 82]	Данные Росстата
	Институциональная среда		Средняя	[45, 53, 77, 103]	
	Банковский сектор		Низкая	[93]	Данные портала Банки.ру
	Венчурный капитал		Низкая		
	Патентная активность		Низкая	[81]	Данные Роспатента
	Затраты на НИОКР		Средняя	[120, 144, 143]	ЕГИСУ НИОКТР
	Налоговый режим		Низкая	[55, 75]	
	Политические факторы		Низкая	[10, 77]	
	Специальные экономические зоны		Высокая	[10, 52, 63, 158]	Данные Министерства экономического развития России
	Защита РИД		Низкая	[93, 135]	Данные Роспатента
<i>Взаиморазмещение бизнеса</i>	Поставщики и покупатели продукции	Формирование синергетического эффекта, снижение издержек, доступ к рынкам	Средняя	[64, 139]	–
	Конкурентная среда		Средняя	[32, 155]	Данные Спарк-Интерфакс
	Диверсификация бизнеса		Средняя	[64, 120]	Данные Спарк-Интерфакс
	Наличие кластеров		Высокая	[14, 131]	Данные Российской кластерной обсерватории

	Деятельность ТНК		Средняя	[36, 47, 66, 155]	Данные Спарк-Интерфакс
	Положение по отношению/внутри агломерации		Высокая	[64, 79, 100]	Данные Open Street Map
<i>Прочие факторы</i>	Природные риски	Оценка рисков при принятии решений о размещении дорогостоящих производственных мощностей	Низкая	[77]	
	Экологическое состояние территории		Низкая	[57, 80, 154]	
	Климатические условия	Фактор при выборе места жительства для высококвалифицированных кадров	Средняя	[94]	WorldClim
	Приграничное/прибрежное положение	Получение преференций от межгосударственного о/межкультурного сотрудничества, локализации экономической активности и развитой инфраструктуры	Средняя	[125, 126, 154]	Данные Open Street Map

Источник: составлено автором.

Научная инфраструктура. Как было отмечено ранее, научная инфраструктура – один из основных факторов размещения ВТП, рассматриваемый в статьях [28]. Университеты и исследовательские центры являются поставщиками кадров и идей для бизнеса, между ними происходят неявные перетоки знаний [70]. Университетские кампусы выступают заметными ядрами пространственной концентрации ВТП. Подчеркивается [28], что важными факторами размещения становятся не только ведущие вузы и институты, выполняющие фундаментальные исследования, но и рядовые практикоориентированные университеты.

Ключевую роль для размещения отдельных сегментов ВТП играют исследовательские центры. В статье, посвященной анализу факторов размещения российского фармбизнеса [17], авторами особо подчеркивается влияние профильных исследовательских институтов, которые, с одной стороны, являются производителями новых технологий и непосредственными участниками научно-

производственного процесса, а с другой – на фоне дефицита кадров поставляют профильных специалистов.

В открытых источниках (мониторингах, рейтингах) доступен широкий набор данных о российских учреждениях высшего образования, включая информацию о количестве и специализации выпускников. Фактор исследовательской функции институтов и вузов можно оценить с помощью данных по участию в НИР.

Человеческий капитал и качество жизни. Важнейшим элементом выбора локации является комплекс факторов, формирующих уровень качества жизни на территории [90]. Хотя это понятия с точки зрения социологии можно трактовать по-разному, в рамках данной работы подразумеваются социально-культурные элементы среды.

Заработная плата – важная детерминанта при выборе локации. Отмечается, что этот параметр неоднороден для разных отраслей – например, сервисный бизнес менее чувствителен к оплате труда работников, в то время как производства, наоборот, ставят более низкий объем оплаты труда как одно из важнейших преимуществ территории [95].

Исходя из имеющихся данных можно выстроить сразу несколько индикаторов оплаты труда. Во-первых, на уровне муниципалитетов на основе официальной статистики, во-вторых, на микроуровне на основе данных корпоративной отчетности.

Фактор стоимости жизни имеет нелинейную природу по отношению к размещению ВТП: на фоне более высоких зарплат квалифицированный персонал, нанимаемый ВТП, до определенного предела готов брать на себя повышенные издержки на проживание, обусловленные более высоким уровнем жизни.

Важным элементом мобильности квалифицированных рабочих является качество школьного образования, доступность медицинской помощи, иногда учитывается развитие рекреационной инфраструктуры [94]. Вопросы криминогенной обстановки редко анализируются в работах, посвященных анализу размещения ВТП, но могут быть актуальны для других отраслей [90].

Производственная и логистическая инфраструктура. Инфраструктурный фактор входит в периметр анализа большинства исследований в области размещения бизнеса. Во-первых, рассматривается влияние близости транспортных узлов и артерий – железных и автомобильных дорог, речных и морских портов, аэропортов.

Во-вторых, фактором в рамках группы является уровень проникновения и развития инженерной (водоснабжения, тепло- и энергоснабжения) и производственной инфраструктуры (зданий и сооружений, в том числе складских, станочного парка), а также расположение по отношению к технопаркам, браунфилдам и гринфилдам (т.е. специальным зонам с заранее подготовленной инфраструктурой) [121].

Фактор стоимости недвижимости или земельной ренты также представлен в большинстве работ о размещении бизнеса. Достаточно часто фактор ренты становится ключевым для выбора локации на разных масштабных уровнях – стран, регионов, отдельных городов, улиц и строений. Кроме того, этот фактор является одним из медиаторов размещения бизнеса внутри агломераций (подробнее см. далее рассмотрение группы факторов взаиморазмещения бизнеса).

Размер оборотных расходов, связанных с платежами за использование воды и электроэнергии, как фактор размещения рассматривается в работах, посвященных моноотраслевым исследованиям, которые связаны с крупными и ресурсоемкими производствами. Для основной части ВТП его влияние на территории России сравнительно однородно.

Данную группу факторов можно оцифровать благодаря открытым данным картографических порталов (в части автомобильных и железных дорог) и различных реестров и регистров (расположение портов и аэропортов), а также открытых данных министерств и ведомств (размещение технопарков). В связи с недоступностью данных аппроксимировать стоимость недвижимости возможно через близость к крупным агломерациям.

Средовые факторы. В первую очередь к этой группе факторов относится общее социально-экономическое развитие территории, т.е. состояние параметров

экономического роста и качества жизни населения: обычно для оценки фактора используют показатель ВВП или ВРП.

Банковский и венчурный капитал часто рассматриваются в работах, посвященных размещению инновационного и/или высокотехнологичного бизнеса [93]. Отмечается, что в некоторых случаях инвестор может принимать очное участие в жизни бизнеса, в связи с чем актуальна его территориальная близость к предприятию.

Важными для ВТП являются аспекты, связанные с производством нового знания и защитой интеллектуальных прав [93]. Затраты на НИОКР являются одним из прокси-индикаторов силы перетока знаний со стороны вузов и исследовательских бюро, а также интеллектуального потенциала территории, квалификации рабочей силы. Защита интеллектуальной собственности обычно рассматривается на уровне отдельных стран, однако есть примеры оценки влияния патентных офисов на размещение бизнеса.

Среди ключевых факторов выбора локации в работах отмечается налоговая нагрузка. Если в России налоговая среда территориально сравнительно гомогенна (за исключением специальных экономических зон), то, например, в США налоговая нагрузка серьезно отличается от штата к штату. При этом возможна нелинейная связь между налоговой ставкой и влиянием на выбор локации в случае перетока налогов на социальные нужды территории, обеспечивающие ее большую привлекательность [75]. Оценка фактора налоговой нагрузки особенно распространена в работах по анализу размещения ТНК, однако также важна и для ВТП – особенно на ранних стадиях его жизненного цикла.

Индикатором влияния налоговых льгот для России может выступать расположение в специальных экономических зонах (СЭЗ). Анализ влияния СЭЗ и промышленных зон разного правового статуса (в том числе «браунфилдов» и «гринфилдов») на выбор локации бизнеса (включая ВТП) является важной темой в современных статьях [10] (в том числе за счет упомянутого выше фактора налогов, феномена перетока знаний).

Фактору политической стабильности также уделяется большое внимание в работах на тему выбора локации. В первую очередь политические условия рассматриваются в контексте иностранных инвестиций, т.е. размещения иностранных предприятий или предприятий с участием иностранного капитала. Так, подчеркивается [77], что высокий уровень политической и макроэкономической нестабильности на национальном уровне является серьезным негативным параметром при формировании стратегии размещения. Однако предприятия с большим опытом выхода на иностранные рынки или фирмы с государственным участием готовы брать дополнительные риски и адаптироваться к нестабильным политическим условиям в стране входа.

Оценить вклад фактора политической среды на уровне предприятий или муниципалитетов весьма трудно без использования опросов и рейтингов. В качестве прокси-индикатора влияния политической среды может использоваться форма собственности предприятия – исходя из обзора литературы, предполагается, что иностранные компании будут располагаться в наиболее благоприятных с социально-политической точки зрения территориях (крупных агломерациях).

Взаиморазмещение бизнеса. В литературе отмечается, что на размещение бизнеса влияют не только контекстуальные объекты, но и другие предприятия – поставщики и покупатели продукции, конкуренты, кластеры и т.д. В свою очередь, исходя из подходов бизнес-демографии, размещение предприятия в дальнейшем само по себе изменяет окружающий экономический ландшафт [133].

Взаиморазмещение производителя и поставщика продукции – классический фактор производства, который анализировался в первых работах по тематике размещения. Несмотря на это, в современных реалиях развития логистических и информационных коммуникаций часто прослеживается тезис о так называемой «смерти географии» [127], т.е. радикальном снижении влияния близости рынков сбыта. Для ВТП вопрос транспортировки продукции действительно может быть вторичен, благодаря ее высокой добавленной стоимости и невысоким

логистическим затратам в структуре себестоимости (в случае с малотоннажными производствами).

С другой стороны, отмечается [131], что колокация предприятий в рамках одной или нескольких отраслей (включая конкурирующие или кооперирующиеся) приводит к появлению формальных и неформальных связей (в том числе в формате кластеров), перетоков знаний и других синергетических эффектов, обусловленных в первую очередь географической близостью. Особая роль отводится неявным эффектам размещения транснациональных корпораций, в особенности с позиции трансфера технологий и компетенций.

Одной из центральных тематик в рамках вопросов размещения является влияние агломерационного фактора на стратегию выбора локации [100, 119]. Агломерационный фактор выступает одним из ключевых вопросов анализа новой экономической географии: учеными этой школы выделяются центростремительные (доступ к рынку и рынку труда) и центробежные (стоимость недвижимости) силы со стороны агломерации. Вследствие этого характер влияния агломерационного фактора имеет нелинейную природу с наличием локальных максимумов в разных поясах агломераций.

Квантификация факторов размещения описанной группы в случае анализа территории России является сложной задачей и ограничивается доступными наборами данных. При сплошном анализе индустрии сформировать связи «потребитель - поставщик» продукции представляется сложно осуществимой задачей, так же как и выявить кооперационные и конкурентные связи фирм. Однако на основе открытых данных можно сформировать перечень формализованных кластеров, оценить размещение бизнеса по отношению к агломерациям.

Прочие факторы. Влияние фактора экологического состояния территории по-разному оценивается в литературе: в основном, данный фактор учитывается бизнесом в последнюю очередь. Так, в работе [45] подчеркивается, что библиографический анализ ранних работ, изучающих влияние экологического

фактора на размещение, выявил его несущественное влияние в независимости от отрасли или территории, рассматриваемых в исследованиях.

С другой стороны, предприятие само может быть источником выбросов, тогда экологический фактор необходимо учитывать с позиций технико-экономического анализа, принимать во внимание правовые ограничения и социальные риски.

Важным при выборе локации представляется фактор природных рисков: в статье “Earthquake risk assessment and optimal risk management strategies for hi-tech fabs in Taiwan” [84] приводится утверждение о необходимости учета сейсмической активности при размещении дорогостоящей производственной инфраструктуры микроэлектроники с целью обеспечения защиты инвестиций и нивелирования страховых рисков.

В случае с анализом российской действительности представляется, что фактор опасных природных явлений выступает критическим только для отдельных территорий – в особенности районов Дальнего Востока и Южной Сибири (землетрясения и наводнения), Северного Кавказа (сели, лавины, землетрясения). Концентрация высокотехнологичного бизнеса здесь невелика и приурочена к наиболее крупным городам, где влияние природно-климатических аномалий максимально нивелируется.

Однако глобальное изменение климата в части увеличения частоты опасных природных явлений в среднесрочной перспективе усилит значимость рассмотрения климатического фактора в стратегическом выборе локации, что потребует формирования единой базы климатических рисков локального уровня.

Одним из рассматриваемых физико-географических факторов размещения является климат территории с позиции влияния на качество жизни людей [87]. В отличие от факторов риска лонгитюдные данные о климатических параметрах территории высокой точности доступны на территории всего земного шара.

Свою роль в размещении высокотехнологичного бизнеса может также играть фактор приграничного/прибрежного положения. Исследования [118, 119] показывают повышенный инновационный потенциал приграничья, связанный с

неявными перетоками знаний, удобной логистикой и притяжением иностранных инвестиций.

1.4 Изменения в целеполагании бизнеса в России и за рубежом

Многие факторы размещения, представленные выше, применимы для анализа размещения высокотехнологичного бизнеса современной России, однако ввиду историко-географических особенностей развития у высокотехнологичного комплекса имеется своя специфика выбора локации.

Так, одну из ключевых ролей в размещении высокотехнологичного бизнеса России играет фактор промышленной базы («эффект колеи») [2]. В свою очередь основная масса крупных промышленных площадок была заложена в период СССР (в том числе вследствие эвакуации промышленности в период Великой Отечественной войны), когда при размещении производства руководствовались нерыночными факторами, в том числе обусловленными стратегической безопасностью. Во многом в дальнейшем это привело к деконцентрации производства и формированию локальных точек размещения.

Практическое развитие концепций ТПК и ЭПЦ в советский период позволило сформировать производственные кластеры с четким разделением функций между их элементами. Одной из задач формирования ТПК было освоение слабозаселенных территорий СССР с высоким ресурсным потенциалом даже при низкой транспортной доступности и сложных природно-климатическими условиях. В рамках промышленных комплексов разных масштабов формировалась сеть НИИ, конструкторских бюро, образовательных учреждений.

Однако в результате процесса перехода к рыночной экономике зачастую происходило разрушение кооперационных связей, приводящее к ликвидации предприятий и утрате промышленного потенциала. Пострадали непрофильные активы социальной инфраструктуры, деградация которых привела к утрате

привлекательности ряда территорий для высококвалифицированных специалистов и резкой депопуляции (особенно в моногородах).

Разворот мировой экономики в сторону сферы услуг и производства нематериальных ценностей серьезно меняет мотивы размещения бизнеса: на первый план выходят не производственные возможности территории, а параметры качества жизни [146], доступ к наиболее квалифицированным кадрам. Подобный тренд в России отчетливо прослеживается на примере крупнейших агломераций, где на месте промышленных площадок (в том числе выпускавших высокотехнологичную продукцию авиаракетостроения, радиоэлектроники и т.д.) были сформированы жилые и офисные пространства, в которых быстрыми темпами развивается сервисный высокотехнологичный бизнес.

Неравномерность социально-экономического развития, условий ведения бизнеса и объемов поддержки предпринимательской инициативы, крайне высокая степень локализации центров принятия решений (как государственных, так и коммерческих) в России привели к чрезмерной концентрации высокотехнологичного бизнеса в Москве [20]. Такая ситуация приводит к оттоку наиболее квалифицированных кадров из регионов и дальнейшему нарастанию диспропорций территориального развития.

Эволюционные процессы в формировании промышленного комплекса высоких технологии ярко прослеживаются на примере Москвы и Московской агломерации – крупнейших в России центров высокотехнологичных производств и услуг. Концентрация финансовых и человеческих ресурсов позволила Москве развивать современные высокотехнологичные предприятия, способные быть конкурентоспособными как на внутреннем, так и на внешнем рынках. Промышленный комплекс территории имеет долгую историю со своими взлетами и падениями, повлиявшими на его территориальную структуру [25].

Первый шаг в промышленном развитии Москвы был сделан в XVIII–XIX вв. с внедрением новых принципов производства: уникальные характеристики физической и экономической географии столицы (доступ к водным ресурсам, узловое расположение на торговых путях) помогли сформировать развитую

текстильную промышленность на периферии города. С бурным ростом Москвы в конце XIX – начале XX в. фабрики, некогда бывшие периферийными, теперь вошли в центральную часть города (среди них хорошо известные Трехгорная и Даниловская мануфактуры, здания которых до сих пор используются в качестве лофт-пространств). К началу Первой мировой войны в Москве сосредоточилось более 25% всего промышленного производства Российской империи.

Процессы индустриализации в СССР начались с Московского региона. За короткий период в 1930-х гг. в городе появились совершенно новые отрасли промышленности: химия, металлургия, автомобилестроение, авиастроение и т.д. С началом Второй мировой войны некоторые из наиболее ценных предприятий Москвы были эвакуированы в Поволжье и на Урал: одни вернулись с окончанием войны, другие – сформировали новые производственные и научные площадки в регионах.

В 1950–1960-е гг. внутри города начал формироваться научно-исследовательский комплекс с новыми конструкторскими бюро, научными центрами и другими элементами научной инфраструктуры. Развитие науки позволило ускорить модернизацию существующих предприятий по производству высокотехнологичной продукции, сформировать совершенно новые передовые ядерные и ракетные научно-производственные комплексы. Рост промышленной мощи Москвы спровоцировал строительство высокотехнологичных предприятий в Московском регионе с целью расширения его концентрации и формирования центров специализации (Королев и Жуковский – аэрокосмическая промышленность, Троицк и Дубна – ядерная наука).

Однако активное промышленное развитие города сформировало совершенно новые проблемы планирования Москвы [4]: нехватка низкоквалифицированных рабочих, экологические проблемы. Рекомендации по размещению промышленности (формирование совместных пространственно ограниченных производств наподобие современных кластеров) и их переносу за пределы города не были приняты во внимание из-за лоббизма в промышленной и научной элитах, которые защищали уже сформировавшиеся кооперационные

связи. Тем не менее до 1990-х гг. Москва по-прежнему оставалась крупным промышленным производителем России.

С распадом СССР московский промышленный комплекс пережил резкий спад [3,8]. Налаженные связи между поставщиками и потребителями были нарушены, экономические проблемы спровоцировали «утечку мозгов и снижение конкурентоспособности технологических решений московских предприятий. Массовые банкротства предприятий приводили к их переориентации на менее технологичные продукты или полному закрытию производства.

Параллельно с этим свой стремительный рост начал сектор услуг Москвы: освободившиеся производственные площади заняли офисы, склады, селитебные территории. Открывшийся потребительский рынок Москвы стал магнитом для иностранных инвесторов, в том числе тех, которые помогли сформировать новый высокотехнологичный бизнес, уже представленный в западном мире [5]: ИКТ, услуги по программированию и т.д. Эти преобразования полностью перестроили экономику Москвы: в 1997 г. около 58% валового регионального продукта города приходилось на сферу услуг, в 2005 г. – 77,8%.

С 2010 г. в Москве начался обратный разворот в сторону развития промышленности, особенно в сфере высоких технологий: городские власти подготовили набор мер поддержки отраслей с акцентом на продвижение технопарков – особых промышленных территорий с развитой инфраструктурой и налоговыми льготами для резидентов. Реализуются проекты инновационного центра «Сколково» и Московского инновационного кластера, объединяющие научную, деловую и образовательную инфраструктуру. В Московской области сформирована сеть ОЭЗ, в том числе на базе наукоградов – Дубны и Фрязино.

Трансформационные во времени процессы в выборе факторов размещения характерны не только для России и СССР. В США усиленное развитие высокотехнологичного сектора началось в 1950–1960-е гг.: в первую очередь, как и в СССР, оно связано с развитием оборонно-промышленного комплекса, что вызвало формирование центров высокотехнологичной промышленности на

территориях с невысоким экономическим потенциалом, но выгодным стратегическим положением (например, на Среднем Западе).

Для своего времени оборонный комплекс стал наиболее инновационным, сформировал абсолютно новые отрасли промышленности, в том числе ракетостроение, а, как следствие, требовал формирования новых производственных площадок. Их локация для сверхдержав стала весьма трудным выбором: с одной стороны, необходимо было засекретить производства от противника любыми способами, обезопасить население в случае чрезвычайной ситуации, рассредоточить по территории для сохранения хотя бы части мощностей на случай конфликта; с другой, – инновационная продукция, выпускаемая подобными предприятиями, требовала максимально компетентного кадрового состава, большая часть производств нуждалась в развитой логистике [115].

В то время как в США военные компании лишь направлялись государством (в том числе посредством контрактов и заказов), формировали собственную линию развития на основе запросов рынка, в СССР в условиях плановой экономики регуляция отрасли производилась намного жестче, хотя при этом ей и предоставлялось куда больше привилегий. В отличие от советского опыта, в США со временем дорогостоящие производства начали «переезжать» в более адекватные рыночным факторам территории (ближе к экономическим и научным центрам, основным очагам расселения, существующим промышленным базам), что впоследствии снизило проблематику моноотраслевых территорий.

Крайне важную роль в формировании территориальной структуры ВТП в США сыграла система вузов. Начиная с 1930-х гг. на базе ведущих технических университетов активно развиваются электроника, радиоэлектроника, электротехника для гражданских и военных целей, контракты с ОПК позволяют инвестировать в развитие кампусов, поддержку перспективных разработок и проектных коллективов для коммерциализации разработок. Наиболее яркой является история Стэнфордского университета, ставшего одним из ядер сформировавшейся Кремниевой долины, в том числе благодаря созданию

якорного промышленного парка, предлагающего экономические и инфраструктурные преференции.

Согласно докладам Еврокомиссии [124] глобальные изменения в целеполагании бизнеса характерны и для Европейского Союза. На смену гомогенизации промышленного пространства ЕС в 1970-х гг. уже начиная с 1980-х гг. наблюдается усиление специализации стран объединения, что сформировало две разнонаправленные тенденции.

В первую очередь, произошла концентрация низкотехнологичного бизнеса на фоне отмирания наименее экономически эффективных предприятий. С другой стороны, для средне- и высокотехнологичных производств, наоборот, наблюдалось формирование новых локализованных производственных центров. Появляющиеся очаги концентрации ВТП характерны для стран с развитым научно-исследовательским комплексом. Эконометрические оценки в рамках доклада указывают на возрастающую роль квалифицированного труда и численности научных работников для размещения производств в целом. Отмечается также нелинейное влияние агломерационного фактора: для производств, использующих полуфабрикаты в производственном процессе, фактор центральности является важным, в то время как для экстенсивных производств (имеющих повышенную ренту от масштаба) важность центральности снижается.

Европейский опыт формирования крупных очагов ВТП Великобритании и Франции рассматривается в статье “Emergence and growth of high-tech activity in Cambridge and Grenoble” [63]. Работа анализирует истории становления ВТП в Кембридже и Гренобле, т.е. на примере территорий с изначально развитой академической средой. В отличие от США форсированное развитие предпринимательства на базе университетов и научных институтов началось только в 1970-х гг., причем в обоих случаях акселератором стало создание опорного научного парка (по примеру крайне успешного опыта Массачусетского и Стэндфордского университетов). Отмечается, что изначальная научная и научно-промышленная специализация академической базы (электроника для

Гренобля и компьютерные вычисления и биотехнологии для Кембриджа) сформировали итоговую специализацию созданных технологических зон.

Для Китая развитие ВТП в исторической перспективе было (и остается) неразрывно связано с иностранными инвестициями, в связи с чем в исследованиях географии высокотехнологичного бизнеса страны большую роль играют вопросы перетока и трансфера знаний, экстерналий от высокотехнологичных предприятий. Кроме того, важным фактором в случае ведения высокотехнологичного бизнеса международными корпорациями является логистика: именно прибрежные районы Китая с выгодным транспортно-географическим положением стали наиболее значимыми для ВТП [154].

Ядрами концентрации ВТП выступают территории с особым режимом: так называемые национальные зоны высокотехнологичного промышленного развития (NHTIDZ – National High Technology Development Zones). Первые зоны были сформированы в начале 1990-х гг., масштабирование проекта произошло уже в 2010-х гг. с нарастанием ввода новых экономических зон в начале пятилетних планов. Отмечается [158], что экономические зоны, помимо прочего, стали драйверами процессов урбанизации, вокруг них выстроилась социальная инфраструктура для привлечения кадров.

Кроме того, необходимо учитывать, что изменения в целеполагании затронули не только страны и отдельные территории, но и отрасли и компании. Это связано с изменениями технологических аспектов производств, логистических возможностей, состояния рынков, управленческих систем и т.д. Наиболее яркими примерами являются крупные японские корпорации, например, «Нинтендо» – компания, начавшая с производства настольных игр впоследствии стала одним из мировых гигантов видеоигровой индустрии, сформировав корпоративную сеть внутри и за пределами Японии.

Выводы по итогам главы. Анализ современной литературы по теме размещения бизнеса в целом и ВТП в частности, а также изучение советского, российского и зарубежного опыта развития высокотехнологичных предприятий

позволяет сформулировать несколько важных для диссертационного исследования тезисов.

Во-первых, прорыв, вызванный увеличением доступности корпоративных микроданных, а также различных открытых данных кратно увеличил возможности количественного анализа влияния факторов размещения и степень его валидности. Исследования показывают широту инструментария, который возможно применить для подробного анализа темы размещения, в том числе методов продвинутой эконометрики и машинного обучения. Однако необходимо учитывать мультимасштабную природу вопроса о размещении: факторы, важные для микроуровня, могут не работать в более мелком масштабе, неоднороден и доступ к данным на разных уровнях приближения.

Во-вторых, в работах указывается на крайне высокую роль научно-образовательной инфраструктуры при выборе локации ВТП как из-за фактора доступа к наиболее квалифицированным кадрам, так и из-за феномена перетока знаний. В связи с этим по ходу исследования был подготовлен большой объем пространственных данных о научной и образовательной деятельности (объем и тематика выполнения НИР в разрезе организаций, количество студентов ВУЗов по направлениям и т.д.) для дальнейшего моделирования.

В-третьих, в современных работах подчеркивается привлекательность территорий с особым статусом (особых экономических зон, научных и технологических парков, формализованных кластеров) для ВТП среди всех стран-лидеров в производстве высокотехнологичной продукции. Это наблюдение актуально и для российского бизнеса, причем оно четко подтверждается еще на верхнеуровневом анализе, представленном далее.

При этом в современных российских исследованиях на тему размещения ВТП и смежных вопросов прослеживается ограниченность территориального охвата (работы ведутся на уровне регионов, либо на локальном уровне в формате анализа конкретных примеров), редко применяются микроданные, способные существенно расширить информационную базу исследований.

2. РАЗМЕЩЕНИЕ ВЫСОКОТЕХНОЛОГИЧНЫХ ПРЕДПРИЯТИЙ НА ТЕРРИТОРИИ РЕГИОНОВ РОССИИ

2.1 Аспекты использования микроданных

По мере интеграции различных акторов экономики (организаций, регулирующих органов, некоммерческих объединений и др.) в интернет-пространство и увеличения их цифрового следа возник спрос на обработку и анализ аспектов деятельности, для изучения которых ранее первичные сведения не были доступны: логистики, рынка труда и научно-технических разработок.

В результате «традиционные» источники статистических данных – готовые отчеты и публикации, перестают удовлетворять растущие запросы различных групп пользователей. Причинами этого являются задержки, связанные с публикационным лагом, и отсутствие возможностей оперативной актуализации методологии сбора и обработки сведений. Более того, формат статистических публикаций вынуждает прибегать к агрегированию на крупном иерархическом уровне, что значительно ограничивает их аналитический и исследовательский потенциал для решения задач размещения производств.

Для обеспечения возможности анализа информации во всех доступных аналитических разрезах многие официальные статистические ведомства обеспечивают возможность доступа к первичным данным (микроданным). В частности, Бюро переписи населения США уже на протяжении нескольких десятилетий предоставляет исследователям анонимизированные индивидуальные записи о состоянии отдельных компаний, а также результаты переписей и обследований населения [58]. Европейская статистическая система, объединяющая Евростат и национальные статистические службы, также обеспечивает доступ к обезличенным итогам опросов и обследований домохозяйств и организаций [46].

Федеральная служба государственной статистики Российской Федерации размещает в открытом доступе микроданные некоторых опросов и исследований,

благодаря чему возможен многомерный анализ различных социально-экономических процессов в разрезе регионов, социальных групп, половозрастной структуры и т.д. Однако аналитические возможности микроданных зачастую ограничиваются их анонимностью, что не позволяет связать их с другими массивами информации и ограничивает их применимость для изучения отдельных специфических аспектов различных отраслей.

Вследствие этого использование альтернативных источников информации, зачастую не являющихся деперсонифицированными, становится необходимостью, а вопросы их обработки, дополнения, агрегирования и визуализации – ключевыми для принятия управленческих решений. В отличие от статистических данных, специально собираемых и предоставляемых для анализа, альтернативные данные зачастую не ориентированы на использование в комплексе, из-за чего требуются их предварительная обработка, верификация и гармонизация.

Так, на бразильском опыте [108] исследованы взаимосвязь краткосрочных индикаторов, составленных на основе Google Trends и Google Mobility, с уровнем инфляции, доходами и другими экономическими показателями. В практической деятельности широко используются данные дистанционного зондирования (спутниковые снимки, аэрофотосъемка, радиолокация и др.) для изучения состояния различных отраслей экономики. Была выявлена тесная взаимосвязь изменения характера застройки [107], растительного покрова, концентрации загрязняющих веществ [63], ночного освещения [104] с экономической активностью на различных иерархических уровнях.

Помимо косвенных показателей развития отраслей экономики на основе альтернативных источников информации большое внимание в исследовательской практике уделяется микроданным, представляющим сведения о каждой единице отбора (индивид, домохозяйство, организация и др.) без агрегирования. Микроданными служат как первичные результаты опросов и исследований, так и записи в базах, реестрах и регистрах, которые также не всегда ориентированы на использование в аналитических целях.

Применение микроданных для исследования социально-экономических процессов и производственной сферы как в России [98], так и в зарубежной практике [54] уже давно получило широкое распространение. Возможности машинной обработки позволили интегрировать в анализ состояние отраслей экономики данные финансово-экономической отчетности предприятий, на текущем этапе широко обсуждаются перспективные направления их использования и потенциальные источники проблем (Таблица 3), связанных с качеством и репрезентативностью [122], а также с необходимостью защиты персональных данных [65].

Таблица 3 – Сравнительная характеристика «традиционной» статистики и микроданных как источников информации

Особенность	Традиционная статистика	Микроданные
Уровень агрегирования	Данные агрегированы по статистическим ячейкам определенного размера	Данные индивидуализированы
Источник данных	Сбор сведений официальной статистической службой	Открытые наборы сведений, собираемых различными агентами
Точность и достоверность	Распространитель данных несет ответственность за их точность	Менее точные данные, содержащие случайные пропуски и ошибки
Защита персональной информации	Данные представлены в агрегированной форме, никаких персональных сведений не распространяется	В случае предоставления неанонимизированных индивидуализированных данных необходимо учитывать вопросы приватности
Необходимость предобработки	Данные готовы к использованию	Может потребоваться дополнительная предобработка
Наличие метаданных	Для каждого набора представлены метаданные и методология сбора	Зависят от конкретного источника

Источник: составлено автором.

Сбор микроданных ведется не только официальными статистическими ведомствами, но и другими государственными и некоммерческими учреждениями [69]. Так, помимо результатов переписей и промежуточного статистического учета оценки численности населения могут быть проведены на основе электоральной статистики, данных операторов связи, ритейлеров, количестве выданных лицензий, удостоверений и др. [67]. Дополнительным преимуществом использования набора альтернативных статистических источников для анализа экономики служит потенциал расширения аналитических возможностей за счет связывания между собой различных массивов и источников, комбинирования сведений. Использование микроданных вписывается в актуальную повестку принятия управленческих решений, основанных на данных (data-driven governance).

2.2 Информационная база исследования

Вопрос обеспеченности данными является ключевым для дальнейшей части работы. В этом параграфе приводится методология сбора разных типов данных, необходимых для решения задач размещения ВТП на мезоуровне (уровне отдельных муниципалитетов). Рассмотрение всей территории России накладывает объективные ограничения на информационную базу исследования (закрытость отдельных источников, трудоемкость их обработки и анализа, отсутствие микроданных или муниципального разреза, диспропорции в качестве и количестве информации для разных территорий), из-за чего ряд важных факторов размещения (например, стоимость недвижимости) был выведен за периметр анализа.

Под высокотехнологичными предприятиями в работе подразумеваются предприятия с основными ОКВЭД2 – 21 и 26, 20, 27, 28, 29 [76,110] (т.е. в периметр анализа на уровне второго знака ОКВЭД2 включены среднетехнологичные предприятия высокого уровня). Подобный подход к формированию периметра распространен в российских и зарубежных работах,

хотя и имеет ряд недостатков и альтернатив. В первую очередь, сам по себе вид деятельности не говорит о степени технологичности предприятия: например, в фармацевтическую отрасль входят предприятия, которые занимаются производством простейших лекарственных средств. Однако в случае со сплошным анализом отрасли влияние подобного отклонения снижается.

Альтернативными методами выявления ВТП выступают дополнительные опросные [92] и статистические обследования, использование статистики выпуска высокотехнологичных товаров. Очевидные проблемы есть и у этих методов: опросы могут покрыть только ограниченную выборку компаний; дискуссионными являются перечни высокотехнологичных товаров, которые формируются как в России, так и за рубежом, а данные по выпуску на уровне предприятий обычно являются коммерческой тайной и не разглашаются статистикой.

Анализ актуального размещения высокотехнологичных производств в рамках исследования основывается на данных системы СПАРК-Интерфакс, агрегирующей данные отчетности предприятий в России и ряде стран СНГ. СПАРК основывается на данных бухгалтерской и статистической отчетности, информации Федеральной налоговой службы, Казначейства, Верховного Суда и др. Благодаря системе можно создавать выборки предприятий по разным параметрам, в том числе основному ОКВЭД 2 (виду деятельности), что позволяет сформировать общую базу данных высокотехнологичных производств по России.

Наличие адреса регистрации предприятия в системе СПАРК позволило получить его координаты с помощью техники геокодирования. В данной работе геокодирование проводилось с использованием языка программирования Python в среде Spyder на основе библиотеки geouru. Геокодирование производилось при помощи API Яндекс.Геокодера.

Использование адреса регистрации в качестве индикатора размещения предприятия является одним из ограничений при анализе. Зарегистрированное в Москве предприятие может вести свою деятельность и в других регионах, иметь несколько производственных площадок. Однако подобная информация не

отслеживается российскими государственными службами и на микроуровне практически безальтернативна (за исключением ручного сбора данных, невозможного при работе в масштабе всей страны).

Геокодирование также может быть фактором смещения данных, однако, учитывая высокий уровень стандартизации адресов компаний в системе СПАРК, можно предположить, что потенциальная ошибка не превышает 5% как в аналогичных работах [96].



Рисунок 4 – Результат геокодирования данных о высокотехнологичных предприятиях

Источник: составлено автором на основе данных БД СПАРК.

Дополнительно из разных источников (на основании анализа, проведенного в параграфе 1.3) были получены данные о факторах размещения бизнеса. Большая часть доступных массивов в соответствии с концепцией открытых данных уже была представлена в минимально предобработанном виде с возможностью выгрузки в машиночитаемом формате.

Однако некоторые источники не обеспечивали прямой доступ к первичной базе данных либо не предоставляли возможности полной выгрузки всей информации. Для получения данных из них на языке программирования Python

были реализованы скрипты, обеспечивающие сбор информации в автоматизированном виде при помощи веб-скрейпинга, производимого средствами сервиса автоматизации Selenium WebDriver. Массивы данных, полученные из различных источников, характеризовались разными уровнями достоверности, качества данных, актуальности и др.



Рисунок 5 – Результат обработки и геокодирования данных об исполнителях НИР
Источник: составлено автором на основе данных ЕГИСУ НИОКТР.

Например, ключевой источник данных о гражданских НИР в России – ЕГИСУ НИОКТР, содержал ошибки, связанные с получением информации на основе заполненных отчетов пользователей. В частности, в ряде случаев было неверно указано число разрядов в объеме финансирования НИР, из-за чего стоимость многих проектов оказалась завышенной. Для устранения систематических ошибок были проведены проверки полноты данных по отдельным организациям, отчитывающимся о научных разработках. Крупнейшие НИР, стоимость которых превышает 100 млн руб., были просмотрены вручную.



Рисунок 6 – Визуализация пространственных данных о линейных инфраструктурных объектах на примере сети железных дорог

Источник: составлено автором на основе данных OpenStreet Map.

После выгрузки, производимой с применением готовых инструментов, предоставленных администратором данных, или веб-скрейпинга, все массивы информации проходили предобработку, верификацию и гармонизацию. Одной из задач предобработки стало получение координат объектов или иных признаков для их геопривязки, для чего использовались поля-ключи (например, название муниципалитета нахождения) или техника геокодирования.

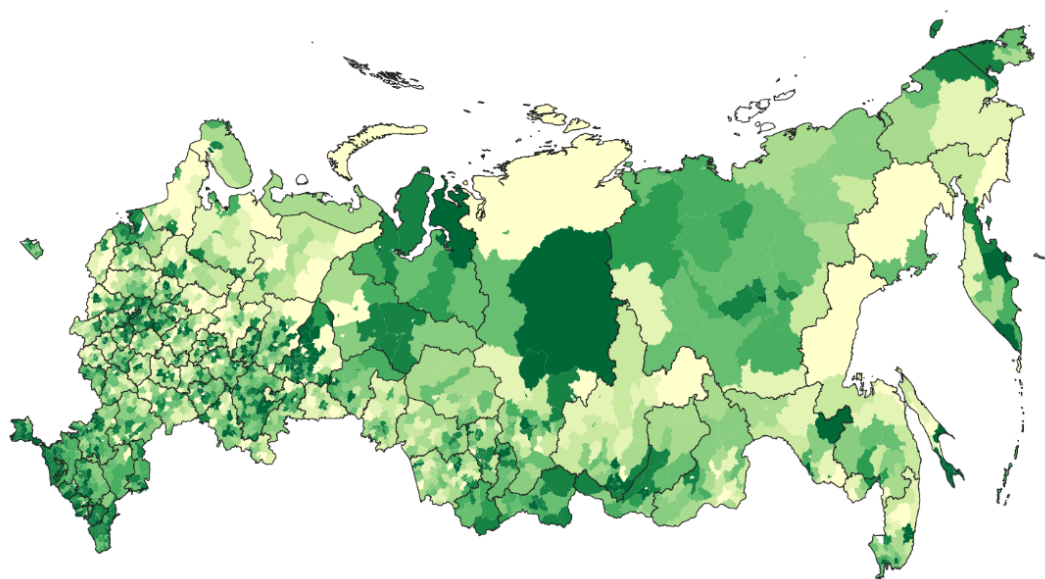


Рисунок 7 – Результат обработки данных о численности населения по муниципальным образованиям

Источник: составлено автором на основе данных БДПМО.

Для большей части объектов (например, исполнителей и заказчиков НИР) геокодирование происходило на основе привязки адреса по ИНН из БД СПАРК-Интерфакс. При отсутствии ИНН проходило менее точное геокодирование на основе всей доступной информации об объекте (например, по названию в сочетании с типом объекта и/или региональной привязкой).

Таблица 4 – Перечень информационных источников

Содержание базы данных	Организация-распространитель данных	Источник данных	Необходим ли веб-скрейпинг	Необходимость предобработки данных
Микроданные предприятий России	Интерфакс	СПАРК-Интерфакс	Нет (доступ через API или выгрузка в .xls)	Нет
Границы стран, регионов,	–	Проект OSM/проект Natural Earth	Нет	Да, фильтрация атрибутов

муниципалитетов; крупных городов				
Данные муниципальной статистики	Росстат	БДПМО	Да (для ускорения сбора данных)	Да (совмещение с сеткой АТД)
Образовательная инфраструктура (вузы)	НИУ ВШЭ	Мониторинг высшего образования НИУ ВШЭ	Да	Да, геокодирование
Образовательная инфраструктура (СПО)	Рособрнадзор	Реестр лицензий на осуществление образовательной деятельности	Нет	Да, геокодирование
Образовательная инфраструктура (школы)	Рейтинговое агентство «Эксперт» (RAEX)	Рейтинг школ России	Нет	Да, геокодирование
Научно-исследовательские разработки	ФГАНУ ЦИТиС	ЕГИСУ НИОКТР	Нет. Выгрузка в формате json	Корректировка данных в ручном режиме
Климатические параметры территории	–	Проект WorldClim	Нет, выгрузка в формате geotiff	–
Социальная инфраструктура (медицинские учреждения)	Росздравнадзор	Реестр разрешений на ведение медицинской деятельности	Нет	Да, геокодирование
Размещение речных портов	Министерство транспорта Российской Федерации	Перечень речных портов	Нет	–
Размещение морских портов	Федеральное агентство морского и речного транспорта	Реестр морских портов	Нет	Да, геокодирование
Размещение аэропортов	Открытые данные	Профильные сайты	Нет	Да, геокодирование
Сеть железных дорог	–	Проект OSM	Нет	Да, дополнение данных и фильтрация атрибутов
Сеть автомобильных дорог	–	Проект OSM	Нет	Да, фильтрация атрибутов
Специальные зоны ведения бизнеса (технопарки, кластеры, промышленные)	Минпромторг России	ГИСП	Нет	Нет

парки)				
ОЭЗ/ТОР/ТОСЭР	–	Открытые реестры	Нет	Да, геокодирование
Банки и банкоматы	–	Открытые реестры и агрегаторы	Да	Да, геокодирование

Источник: составлено автором.

В общем итоге для дальнейшего анализа сформирована база пространственных данных с информацией о более чем 30 показателях, представляющих влияние факторов размещения, релевантных для высокотехнологичной промышленности: всего в базе находится информация о более 100 000 точечных объектах (включая данные о 30 910 высокотехнологичных предприятиях), а также более 2300 площадных и 7000 линейных объектах.

2.3 Анализ размещения ВТП России с применением микроданных

В соответствии с информационным периметром анализа, представленным в параграфе 2.2, было проведено картирование высокотехнологичных предприятий. Для упрощения интерпретации выручка предприятий (как индикатор масштаба деятельности) была агрегирована по сетке муниципальных образований в разрезе отдельных ОКВЭД 2, соответствующих высокотехнологичным производствам.

База геоданных (включающая информацию как о ВТП, так и о факторах размещения) была проанализирована на разных масштабных уровнях (и на общероссийском, и на уровне отдельных муниципалитетов) для формирования качественных выводов об особенностях размещения предприятий. Выявленные типичные закономерности размещения были рассмотрены на примерах конкретных предприятий и/или муниципальных образований в формате примеров.

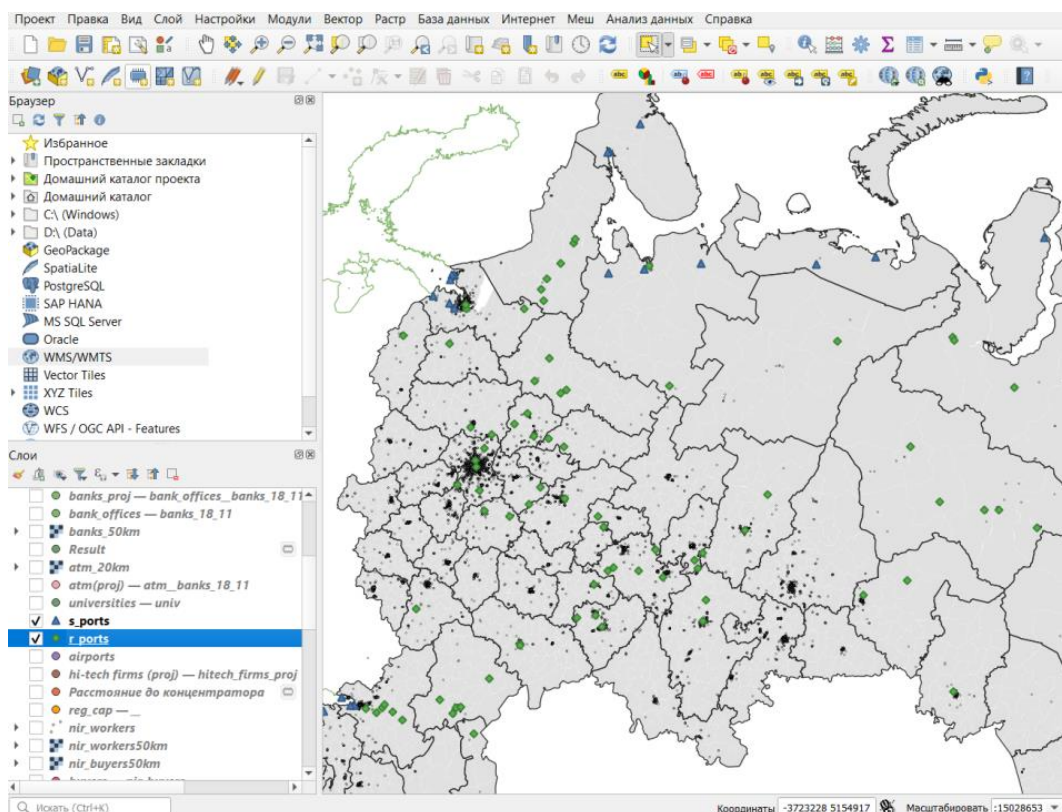


Рисунок 8 – Визуализация рабочей среды исследования с набором пространственных данных

Источник: составлено автором

Для удобства анализа в ходе построения качественных суждений для оперативного сопоставления размещения предприятий и предположительных факторов размещения использовалась среда QGIS (Рисунок 8), а также отдельные модули представленной ранее СППР (в частности, по оценке научной инфраструктуры).

Полученные результаты. Анализ картографических материалов, подготовленных на основе информации БД СПАРК (Рисунок 9), указывает на ряд территориальных особенностей в размещении ВТП.

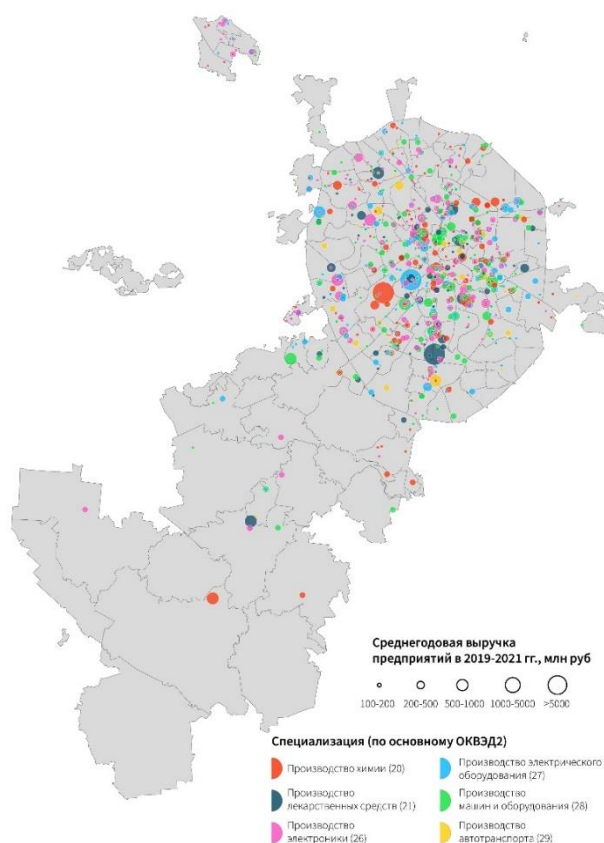


Рисунок 9 – Локация высокотехнологичных предприятий в Москве по состоянию на начало 2022 г.

Источник: составлено автором на основе данных БД СПАРК.

Высокий уровень централизации. Московский регион является абсолютным лидером по числу и объемам выручки ВТП. Для ядра агломерации – Москвы характерны внутригородские особенности размещения высокотехнологичного бизнеса, например, тяготение к центру города (в отличие, например, от примеров Берлина или Парижа); концентрация в поясе НИИ и вузов на западе и юго-западе города, а также наличие субцентров в технопарках.

Стоит также выделить агломерации городов-миллионников (например, Самарско-Тольяттинская – химия, автомобилестроение; Ростовская – точное машиностроение, химия, электротехника; Уфимская – промышленное оборудование, химия и т.д.) как видимые полюса концентрации ВТП.

Примером центра развития высокотехнологичных предприятий в пределах агломерации является Дзержинск – один из общероссийских (а ранее и общесоветских) центров химической промышленности. На сегодняшний день в городе функционируют предприятия по производству крупно- и малотоннажной химии, в том числе специальных полимеров. Большая часть предприятий (как крупных заводов, так и экспериментальных производств) сформированы на основе советских площадок, однако есть и новые производства, нацеленные на создание небольших объемов сложной продукции.

Факт концентрации размещения высокотехнологичных производств в рамках агломерации указывает на актуальность факторов социально-экономического развития. С одной стороны, агломерации являются центром подготовки и притяжения квалифицированных кадров, с другой – комфортной для сотрудников локацией с повышенной транспортной доступностью. В пределах агломераций важными объектами тяготения высокотехнологичных производств являются научные центры, крупные промышленные площадки и специально созданные офисно-производственные комплексы [19].

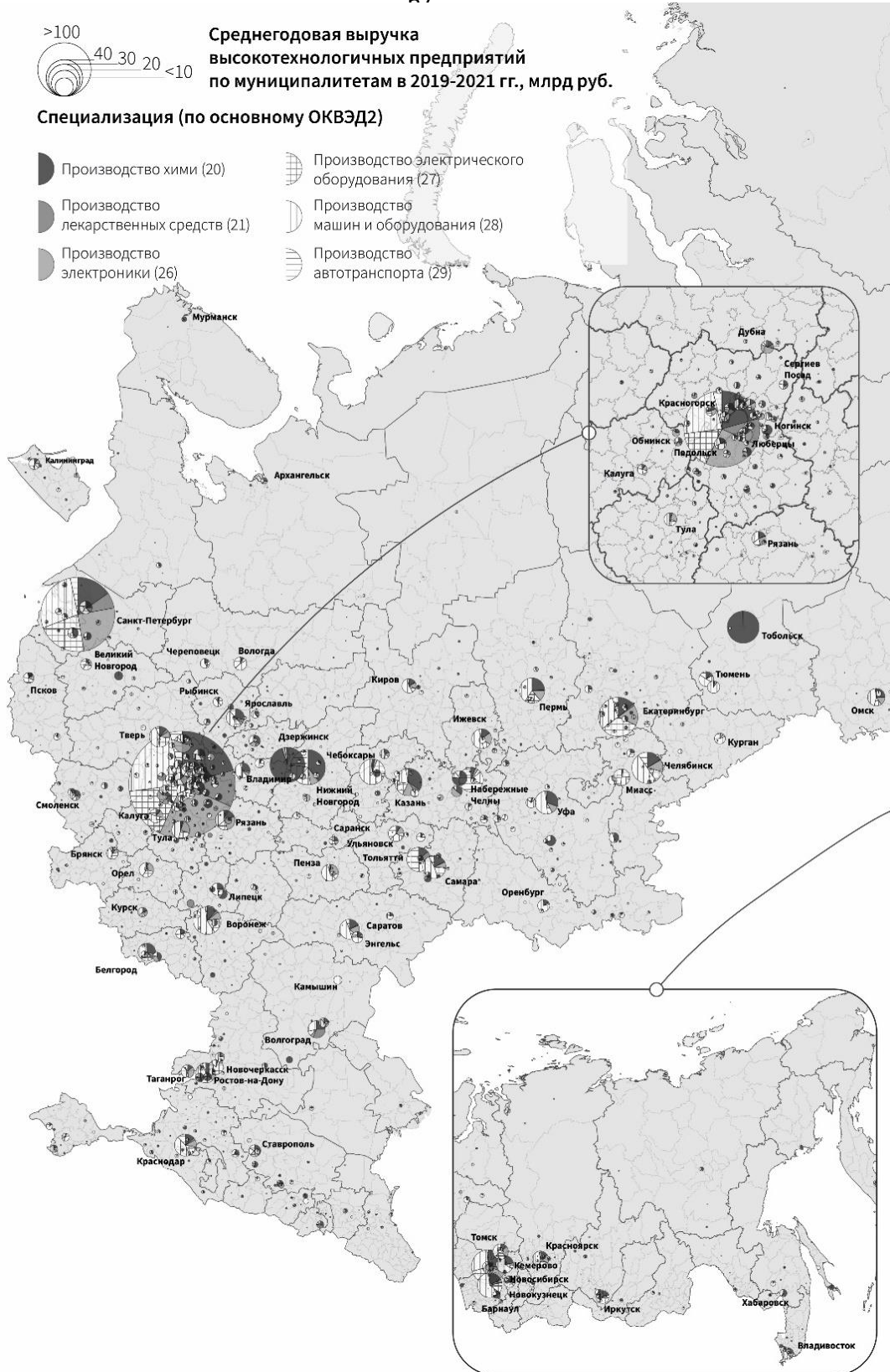


Рисунок 10 – Локация высокотехнологичных предприятий на низовом административном уровне по состоянию на начало 2022 г.

Источник: составлено автором на основе данных БД СПАРК.

Диспропорция Восток – Запад. Очевидна глобальная диспропорция Восток – Запад, характерная для многих социально-экономических явлений в стране. Несмотря на наличие крупных центров развития ВТП за Уралом, распространенность предприятий носит очаговый характер вокруг крупнейших населенных пунктов.

За пределами региональных столиц примерами территорий высокой концентрации ВТП являются Тобольск (нефтехимия), Новокузнецк (горнодобывающее и промышленное оборудование), Бердск (промышленное оборудование), Бийск (фармацевтика и косметика).

Ключевым фактором размещения в данном случае является фактор спроса со стороны более развитых на конкретной территории отраслей – высокотехнологичный бизнес в данном случае выполняет обслуживающую функцию для нефтегазового сектора и горнодобывающей промышленности.

Лакуны в Европейской России. Точечная распространенность ВТП (за исключением региональных столиц) также наблюдается в Нижнем Поволжье (Камышин, Урюпинск, Котельниково) и на Северо-Западе Европейской России (Северодвинск, Череповец, Ухта). Местные предприятия занимаются производством техники для нефтегазовой отрасли, промышленного оборудования, химии на локальные рынки сбыта.

Исключением является Калининградская область – один из центров радиоэлектронной промышленности в эпоху СССР и особая экономическая зона в современной России. Сейчас, помимо самого Калининграда, высокотехнологичные производства находятся в Черняховске и Советске – регион специализируется на производстве электроники, электрооборудования, автомобильной техники. Важным фактором развития высокотехнологичных отраслей в регионе стали кластерные инициативы.

Крупные ареалы размещения в Урало-Поволжье. Урало-Поволжье – вторая по значимости территория распространения ВТП после Московского региона. С одной стороны, это следствие эвакуации предприятий в военное время и крайне развитой промышленной базы, с другой – сравнительно высокого

социально-экономического развития регионов. ВТП здесь размещен как в крупнейших городах и их агломерациях, так и в городах второго порядка (Набережные Челны, Сарапул, Димитровград и др.).

Среди ключевых территорий развития высокотехнологичных предприятий можно отметить Миасс – один из российских центров авиаракетостроения, автомобильной промышленности. Помимо крупных предприятий оборонно-промышленного комплекса, в городе функционирует сеть малых и средних предприятий, обеспечивающих деятельность лидирующих производств. Например, высокотехнологичные МСП вовлечены в кооперацию с Уральским автомобильным заводом, дооснащают выпущенные с завода автомобили под специальную технику.

Примечательно, что важную роль в формировании новых высокотехнологичных производств в районе играют различные виды кластерной кооперации, как неформальной (на примере Миасса), так и формальной – в виде юридических форм и/или технологических и индустриальных парков. Очагами концентрации ВТП также являются особые экономические зоны, например, Алабуга и Тольятти.

Субцентры размещения в Центральной России. Сплошной ареал размещения высокотехнологичного производства в Центральной России выстроен вокруг Московской агломерации и выходит за пределы административных границ Московской области. Наиболее интересным здесь является рассмотрение крупных центров ВТП за пределами столичных агломераций. Анализ показывает наличие городов второго порядка, которые сопоставимы с региональными столицами по масштабам высокотехнологичных производств.

Важными субцентрами являются наукограды – Обнинск, Королев, Жуковский и в первую очередь Дубна. Так, на базе особой экономической зоны Дубна работают предприятия по производству радиотерапевтической и другой медицинской техники, фармацевтической продукции, электротехнического оборудования. В случае Дубны работает синергетический эффект, когда научно-производственные мощности, созданные в советский период, дополняют

современной жилищной и инженерной инфраструктурой, а также создают привлекательные условия для ведения бизнеса.

Выделяются также города второго порядка за пределами Московской области – Рыбинск (промышленное оборудование и продукция для аэрокосмического комплекса), Углич (энергетическое оборудование), Ковров (продукция для ядерного комплекса), Гусь-Хрустальный (промышленная электротехника и нефтегазовое оборудование), Александров (электроника и электротехника). В основном высокотехнологичные предприятия в этих городах – это реорганизованные производства советской эпохи, также необходимо отметить высокую роль оборонно-промышленного комплекса в их развитии.

Анализ ареалов размещения ВТП позволяет выделить отдельные факторы размещения, актуальные для высокотехнологичных предприятий на текущий момент и архетипы стратегического поведения компаний – агломерационную, эволюционную, симбиотическую, сервисную и инкубационную стратегии [19].

В первую очередь очевидно влияние агломерационного фактора и сопряженных с ним средовых факторов, формирующих **агломерационную стратегию** размещения предприятий. Более высокий уровень жизни в крупнейших агломерациях России, наличие там квалифицированных трудовых ресурсов, образовательных и научных площадок, транспортная доступность и развитая система технопарков делают агломерации очевидным преимуществом для высокотехнологичных производств. При этом внутри агломераций на микроуровне есть своя территориальная диверсификация по уровню привлекательности для размещения ВТП.

Рассмотрение крупнейших предприятий в каждом из выделенных ареалов подтверждает влияние «реликтовых» промышленных площадок и производств, сохранившихся после распада СССР. При этом унаследованные современной Россией крупные промышленные производства могут играть разную роль в функционировании высокотехнологичных производств.

Самый простой случай, когда предприятия сохранили свой функционал и приспособились к современным экономическим и технологическим реалиям либо

как единое предприятие, либо как сеть зависимых или независимых производств на основе крупного реструктурированного предприятия. В данном случае стратегию размещения можно охарактеризовать как **эволюционную**.

Такие предприятия могут быть также ядрами для формирования промышленной кооперации, тем самым формируя вокруг себя цепочку поставщиков высокотехнологичной продукции (что особенно явно проявляется в центрах автомобилестроения) в виде неформальных кластеров. Таким образом, наблюдается **симбиотическая стратегия** размещения.

Показательным является размещение высокотехнологичных производств вне основных ареалов размещения: ключевым фактором в данном случае выступает наличие доминирующей отрасли, которая формирует запрос на высокотехнологичную продукцию – в основном это нефтегазовый и топливно-энергетический комплексы. Подобная модель размещения представляется как **сервисная**. Кроме того, наблюдение подтверждает, что фактор взаиморазмещения по отношению к предприятиям разного уровня технологичности может иметь значимость.

Что касается абсолютно новых производств, среди прочего ядрами концентраций для них выступают особые экономические зоны, технологические и индустриальные парки, т.е. локации с дополнительными преференциями в виде готовой инфраструктуры и льготных условий для ведения бизнеса. Ввиду большого срока окупаемости инвестиций для высокотехнологичных производств, их поддержка позволяет снизить смертность бизнеса на начальных этапах его развития, в связи с чем фактор размещения в преференциальных зонах имеет крайне высокий вес при выборе локации. В данном случае стратегия ВТП при выборе локации – **инкубационная**.

Таблица 5 – Результаты качественного анализа размещения высокотехнологичных предприятий

Архетип размещения предприятия	стратегии	Ключевые факторы размещения
<i>Агломерационная</i>		Доступ к кадрам, рынку сбыта, транспортной инфраструктуре и другим преимуществам агломерации
<i>Эволюционная</i>		Развитый промышленный комплекс, унаследованный со времен СССР
<i>Симбиотическая</i>		Близость к крупным промышленным площадкам комплиментарной отрасли
<i>Сервисная</i>		Близость к крупным промышленным предприятиям низко- и среднетехнологичного бизнеса
<i>Инкубационная</i>		Близость к зонам с особым экономическим статусом, одготовленной структурой

Источник: составлено автором.

Качественный анализ размещения высокотехнологичного бизнеса указывает на наличие таких факторов выбора локации, которые, во-первых, подробно не рассмотрены в научных работах (влияние опорной отрасли), а во-вторых, крайне тяжело (или вообще невозможно) оцифровать на уровне нескольких отраслей в масштабе всей страны.

В первую очередь подразумевается критическое влияние фактора реликтовых промышленных площадок. Дробление, реорганизация и переформатирование советских промышленных площадок не позволяет «отследить» их ни по атласам промышленности СССР (к тому же имеющих масштабные и стратегические ограничения), ни по современным базам данных корпоративной отчетности (что связано с почти стопроцентным переоформлением предприятия и невозможностью отследить его возраст без ручного поиска в корпоративных документах).

Кроме того, при переходе к анализу размещения на микроуровне существуют объективные факторы, ограничивающие перечень потенциальных

локаций из-за ограниченности предложения помещений и земельных участков, экологическим и технико-экономическим особенностям производств и т.д. Таким образом, при проведении практикоориентированных изысканий по выбору локации для бизнеса экспертное знание для валидации работы количественных методов и конечного выбора локации является безальтернативным.

Выводы по итогам главы. Развитие концепции открытых данных, распространение микро- и пространственных данных позволяет решать задачи размещения с недоступной ранее пространственной точностью. Оценка находящихся в открытом доступе баз данных показывает, что их содержание закрывает потребность в большей части информации как о высокотехнологичных предприятиях, так и о факторах их размещения.

Качественный анализ размещения высокотехнологичного бизнеса позволил сформировать актуальную картину географии размещения ВТП в России, подчеркнув наличие крупных территориальных разрывов, а также наличие регионов-лидеров (групп регионов и территориальных общностей) по распространению ВТП.

Анализ также показал наличие объективных факторов размещения производств исходя из выбора ими архетипичных стратегий – агломерационной, эволюционной, симбиотической, сервисной или инкубационной. Исходя из типичных для территории стратегий региональные власти могут выстраивать уточненную промышленную и инвестиционную политику, например, стимулировать кластерные инициативы.

Экспертная оценка факторов размещения показала невозможность их полной оцифровки, а также наличие отдельных факторов размещения, определяющих выбор локации уже на уровне агломерации. В связи с этим необходимо учитывать ограничения любых количественных методов выбора локации, подкреплять их экспертизой с пониманием региональной и локальной специфики.

3. МЕТОДИЧЕСКИЕ АСПЕКТЫ РЕШЕНИЯ ЗАДАЧИ РАЗМЕЩЕНИЯ ВЫСОКОТЕХНОЛОГИЧНЫХ ПРЕДПРИЯТИЙ

3.1 Пути решения задач размещения

С точки зрения экономики региона проработка выбора локации решает задачу обоснования долгосрочных рискованных инвестиций на открытие нового производства. Научные и научно-практические работы позволяют обозначить несколько возможных решений задач по выбору локации.

Во-первых, это использование качественных (аналогично параграфу 2.3) и качественно-количественных суждений с позиций отраслевой экспертизы [34]. В этом случае используется экспертное мнение отраслевых специалистов (или совокупность таких мнений с использованием или без использования нормализации либо весовых коэффициентов), основанное на предшествующем опыте размещения производств, а также неявных знаний о преимуществах отдельных факторов размещения, технико-экономических особенностей бизнеса, институциональных особенностей. Очевидным плюсом подхода является простота его интерпретации, однако понятны и риски в виде возможной необъективности экспертной группы.

Распространенным методом повышения качества экспертных суждений является применение аналитических моделей и алгоритмов принятия решений – например, метода анализа иерархии (МАИ). Суть метода в приложении к задачам размещения [156] заключается в формировании иерархической системы факторов размещения и их попарного сравнения на основе экспертного суждения с присвоением весовых коэффициентов для формирования итогового рейтинга потенциальных локаций. Экспертами в решении этой задачи могут выступать не только профессионалы в вопросах размещения, но и потенциальные стейкхолдеры внутри компании, представители власти и/или регуляторов. Очевидно, что для разных групп заинтересованных лиц приоритеты в выборе локации

могут существенно различаться, в таком случае МАИ позволяет прийти к определенному консенсусу.

Во-вторых, использование финансово-экономических и/или эконометрических моделей. В данном случае модели обычно служат для оценки территории с позиции минимизации издержек на размещение или максимизации прибыли. На практике применяется широкий спектр моделей самой разной сложности [42,48,68]: от простых балансов до продвинутой эконометрики. Для задач размещения наиболее распространены являются МНК-регрессии, а также пуассоновская и отрицательная биномиальная регрессии (в связи с аспектами распределения географических явлений), также применяются пространственные регрессии.

В-третьих, все более популярным становится использование моделей машинного обучения [49] для принятия решений о размещении производств. Многообразие и гибкость инструментов машинного обучения позволяют решать самые разные задачи выбора локации. При этом машинное обучение требовательно к качеству и количеству данных, а использование различных видов больших данных (например, данных дистанционного зондирования) устанавливает минимальный порог вычислительных мощностей для моделирования.

При использовании метода «черного ящика» (хотя современные алгоритмы позволяют приближать их к «серому ящику») машинного обучения могут возникать казусы интерпретации моделей и вопросы к их практической применимости. Так, в ряде статей, посвященных сравнительному анализу различных подходов к размещению, результаты решения локационной задачи алгоритмами машинного обучения показали результат, сопоставимый с броском монеты. Однако другие исследования [106] показывают весьма высокий уровень предсказательной способности в сравнении, например, с эконометрическими моделями.

С практической точки зрения в России машинное обучение активно используется для решения задач оптимизации розничных сетей (в частности, в

ритейле и банковской среде). На основе экспертных интервью известно, что крупнейшие банковские организации применяют большие массивы пространственных данных для оценки потенциала отделений банков на основе как их внутренней отчетности, так и близлежащих контекстных объектов, графа транспортной сети и пассажиропотоков, данных сотовых операторов и т.д.

Оптимальным представляется комбинация методов [35,39] для нивелирования их недостатков, однако в таком случае увеличивается объем работы по сбору информации и формированию адекватной методологической конструкции. Однако именно подготовка данных является наиболее времязатратным элементом в ходе задач поиска оптимальной локации, что связано в том числе с использованием пространственных данных, для формирования массива и обработки которых требуются дополнительные компетенции и мощности.

3.2 Интерактивные системы поддержки принятия решений

Для работы с большими массивами данных нужна среда для их обработки и/или визуализации. На практике уже с конца 1980-х гг. [35] в научно-практической среде применяются геоинформационные системы (ГИС) – программные комплексы для хранения, обработки и визуализации пространственных данных. ГИС позволяют обрабатывать данные о факторах размещения, локации компаний, формировать наборы данных с помощью геопривязки и пространственной агрегации, обогащать такие наборы информацией, полученной с помощью инструментов пространственного анализа [86].

Несмотря на то, что аналогичные задачи можно делать напрямую в рабочей среде языков программирования (например, Python или R) благодаря удобным визуальным интерфейсам и сравнительно низкому входному порогу ГИС являются более распространенным решением для работы с пространственными данными, в том числе с последующим экспортом в другие программные

комплексы. При этом ГИС все еще требуют аналитических навыков, интерпретация результатов пространственного анализа или формирования различных индексов и рейтингов на основе пространственных данных требуют специальных компетенций.

Другим путем построения эмпирической базы для исследования в области выбора локации является построение систем поддержки принятия решений на основе VI-платформ [88,128,129,157]. Благодаря описанной ранее цифровизации отраслей экономики и государственных служб возможно проводить более глубокую аналитическую работу, формировать новые инструменты, в том числе интерактивные аналитические панели [29,113,130].

Использование микроданных и интерактивных аналитических панелей (дэшбордов) для их визуализации уже давно вошли в стандартный набор подходов к принятию решений в коммерческом секторе, государственными учреждениями, и институтами развития, в исследовательскую практику. Например, для визуализации базы данных Short-Term Mortality Fluctuations (STFM), содержащей в себе гармонизированные ряды данных демографической статистики по неделям, был разработан интерактивный сервис с возможностью прогнозирования и получения демографических оценок [87]. Применение дэшбордов актуально как при выборе локации для нового предприятия, так и для оптимизации существующей территориальной структуры бизнеса на разных территориальных иерархических уровнях.

Информация из VI-систем представляется в виде удобных визуализаций, для работы с которыми нужны лишь базовые аналитические навыки, что позволяет использовать их при работе с экспертным мнением (например, для визуализации первоначальных гипотез или валидации экспертного суждения). С другой стороны, большинство VI-решений поддерживают возможность выгрузки данных, что позволяет использовать их как библиотеки для использования в последующем моделировании оптимальной локации.

Использование VI-систем в качестве системы поддержки принятия решений (СППР) для визуализации научной инфраструктуры и логистики водным

транспортом в качестве факторов размещения, а также узкоотраслевых факторов размещения фармацевтической промышленности было апробировано в рамках НИР «Разработка интерактивной платформы по оценке состояния ключевых отраслей экономики с применением альтернативных источников статистики».

Первичный этап при создании дэшборда – отбор релевантных источников, производился, в первую очередь, при помощи поиска на порталах открытых данных и официальных веб-сайтах государственных учреждений: министерств и федеральных служб, различных отраслевых ведомств и др. Проводился также поиск сведений на различных некоммерческих ресурсах: научно-исследовательских учреждений, баз открытых данных, интернет-энциклопедий и сервисов со свободным доступом.

Предобработка требовалась для приведения данных, предоставляемых в различных форматах, к единому виду. Статистические публикации и наборы свободно распространяемых данных, как правило, уже были предоставлены в табличной форме (.xlsx, .csv и др.). Однако в некоторых источниках (в частности, база данных научно-исследовательских разработок ЕГИСУ НИОКТР, а также реестр разрешений на ведение медицинской деятельности) исходная структура данных различалась (.json и .xml соответственно). Вся необходимая информация из них также была приведена к табличному виду (формат .csv для последующего обеспечения работы с облачным хранилищем).

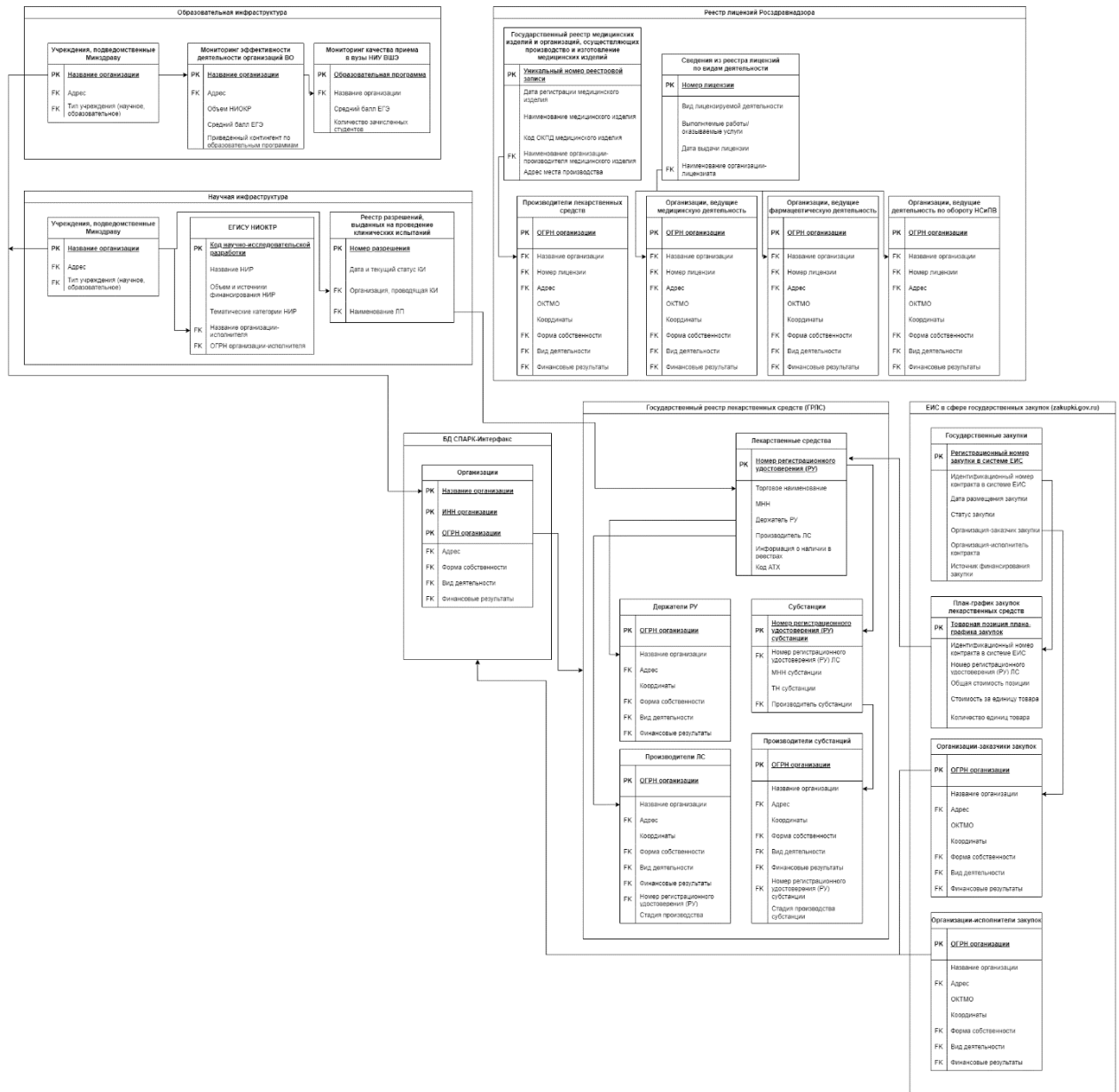


Рисунок 11 – Архитектура данных СПДР для фармацевтической промышленности

Источник: составлено автором.

Все используемые данные были представлены в табличной форме в агрегированном виде. Связь между ними в базе данных производилась с использованием заранее разработанной системы ключей-идентификаторов (Рисунок 11). Все итоговые таблицы, представленные в формате .csv, были загружены на облачное хранилище Yandex.Cloud для формирования базы данных с применением СУБД ClickHouse. В рамках базы данных отдельные модули

(университетское образование, НИР) были реализованы в качестве наборов связанных между собой таблиц.

На этапе сбора информации и поиска возможных решений была проведена сравнительная характеристика основных сервисов интерактивной визуализации данных. Исходя из доступности для освоения, набора аналитических возможностей и потенциала масштабируемости была выбрана платформа Yandex DataLens на основе облачных технологий Yandex Cloud. Интеграция аналитического сервиса с облачными решениями позволяет существенно упростить процесс разработки за счет возможностей использования удаленных хранилищ данных и вычислительных мощностей. В частности, все выгруженные и предобработанные таблицы были размещены на базе виртуальной машины Compute Cloud, обеспечивавшей возможность проведения облачных вычислений.

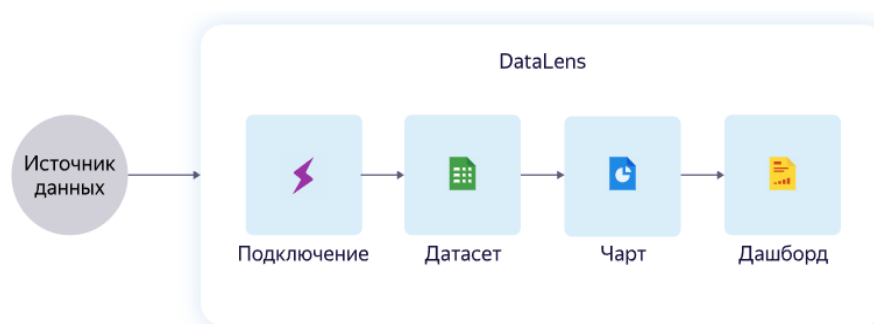


Рисунок 12 – Схема работы DataLens

Источник: документация Yandex DataLens.[160]

Архитектура сервиса Yandex DataLens предполагает использование структуры «Подключение» для осуществления доступа к базам данных на виртуальной машине. Каждое подключение может содержать одну или несколько таблиц с данными (структура «Датасет»). На основе реляционных связей между таблицами также доступно формирование датасетов, содержащих сведения из нескольких таблиц в пределах одного подключения.

В функционал DataLens входят частичная предобработка и агрегирование сведений, получаемых из структуры «Датасет», однако ее возможности

ограничены, из-за чего размещение готовых таблиц на облачном хранилище предварялось их предобработкой и приведения к соответствующей форме для конечных визуализаций. На основе таблиц с данными создаются визуальные представления (структура «Чарты»). В рамках подготовки чартов также возможны донастройка и изменение типа визуализируемой информации (формирование рейтингов, обработка формата записи, приведение к определенному типу данных и т.д.).

Конечный продукт (структура «Дашборд») состоит из набора чартов, размещенных на нескольких вкладках. На каждой вкладке дашборда при помощи фильтров и селекторов существует возможность создания выборки в любом из доступных аналитических разрезов, также в пределах одной вкладки обеспечена возможность установления взаимосвязей между разными элементами. Наконец, инструменты DataLens позволяют дополнить визуализации текстовыми описаниями и графикой.



Рисунок 13 – Визуальная интерпретация системы принятия решений на примере данных о логистике водным транспортом

Источник: составлено автором.

Вкладка с данными о логистике водным транспортом (Рисунок 13) содержит преобработанные данные российского Речного регистра, а также сведения о морских портах с оценками пассажиро- и грузопотоков, в том числе по типам грузов.

На Рисунке 14 показан образ результата визуализации факторов размещения научной инфраструктуры. На вкладке «Наука и образование» приведены данные ЕГИСУ НИОКТР и Мониторинга высшего образования НИУ ВШЭ. Наличие сведений на уровне организаций позволило произвести агрегирование всех данных по регионам и муниципалитетам. Из результатов Мониторинга высшего образования за 2018–2022 гг. представлены средние баллы ЕГЭ абитуриентов в разрезе регионов и специальностей с возможностью детализации информации по университетам. Аналогично приведены сведения о количестве поступивших на ту или иную специальность. Данные о научно-исследовательских разработках, начатых в 2016–2022 гг. ЕГИСУ НИОКТР, агрегированы по организациям-исполнителям. На основе тематических рубрик отдельно были выделены НИР, относящиеся к фармацевтике. На карте представлены затраты на НИОКР по муниципалитетам, также по крупнейшим организациям-исполнителям приведены сведения о НИР по годам.

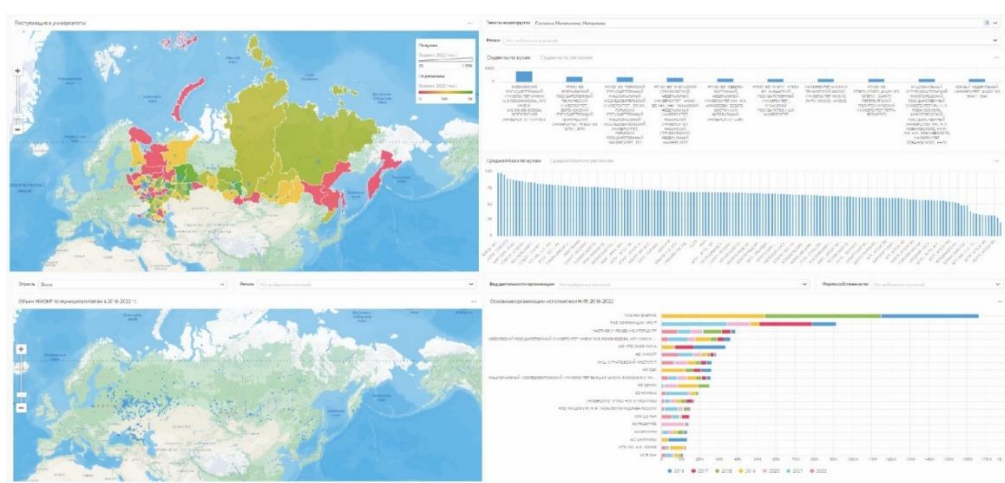


Рисунок 14 – Визуальная интерпретация системы принятия решений на примере данных о научной инфраструктуре

Источник: составлено автором.

На сформированной отраслевой вкладке «Фармацевтика» (Рисунок 15) продемонстрированы данные о факторах размещения, специфичных для фармацевтической отрасли, в том числе о государственных закупках лекарственных средств, данных размещения медицинских учреждений, сведений о заболеваемости и смертности, полученных на основе официальной статистики. Информация о государственных закупках показана в агрегированном виде по муниципалитетам организаций-заказчиков закупок (это позволяет продемонстрировать потенциальный фактор спроса на определенной территории) с возможностью выбора отдельных групп в соответствии с анатомо-терапевтической классификацией лекарств (т.е. отдельный сегмент рынка).

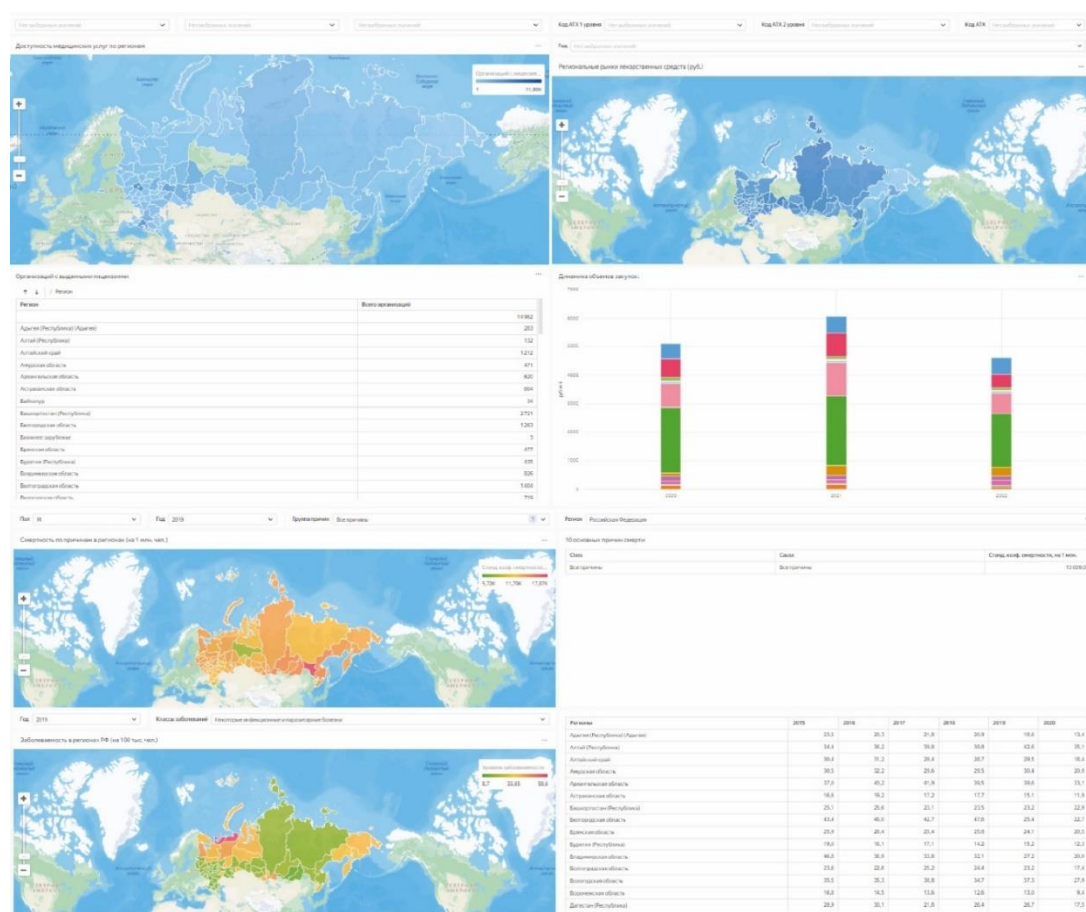


Рисунок 15 – Визуальная интерпретация системы принятия решений на примере данных о фармацевтическом рынке

Источник: составлено автором.

Реестры выданных лицензий на ведение того или иного вида медицинской деятельности использованы для визуализации обеспеченности определенными видами медицинских услуг по регионам. Данные Минздрава и ЦДИ РЭШ о заболеваемости и смертности соответственно приведены по годам и регионам для демонстрации эпидемиологической ситуации в России.

Представленный путь формирования отдельных модулей СППР при размещении высокотехнологичных производств воспроизводим, подходит для визуализации и анализа других высокотехнологичных отраслей и факторов их размещения. СППР применима для построения и валидации экспертных гипотез, возможна ее интеграция в бизнес-процессы организации, исследования. СППР может служить базой данных для машинного обучения, а также для визуализации его результатов и донастройки информационной базы для улучшения качества прогностических моделей. Помимо открытых данных, оператор СППР может интегрировать в систему конфиденциальную информацию, полученную из внешних источников или подготовленную самим оператором (например, оценки рыночных перспектив, данные о конкурентной среде, институциональных особенностях территории и т.д.).

Формирование соответствующего инструмента ложится в логику развития государственных информационных систем (в том числе на основе пространственных данных), а также федеральных и региональных инвестиционных порталов. Подготовка VI-инструмента для оценки потенциала территорий для размещения высокотехнологичных предприятий может позволить региональным органам власти проводить настройку инвестиционной политики на территории, упростит обоснование предложений по размещению технологических и индустриальных парков, зон с особым экономическим режимом.

3.3 Использование машинного обучения для поиска благоприятной локации для ВТП

Машинное обучение для решения задач, связанных с пространственными явлениями, в первую очередь особенно активно используется в приложении к разным направлениям физической географии, климатологии, геоэкологии. Благодаря моделям машинного обучения формируются современные метеорологические и климатические прогнозы, строятся геологические и геоморфологические профили, решаются задачи выделения ареалов отдельных видов флоры и фауны, географического распространения болезней (в рамках медицинской географии).

Наиболее коммерчески успешным стало направление использования машинного обучения в рамках агрогеографии при классификации результатов дистанционного зондирования для определения вегетативного состояния и типа посевов, а также характеристик почвенного покрова и потенциальной урожайности культур.

В социально-экономической географии и регионалистике, пространственной экономике инструменты машинного обучения имеют заметно менее обширное распространение, что в первую очередь связано со сложной природой социально-экономических явлений, а также ограничениями, связанными с доступностью объективных микроданных. Достаточно популярной тематикой является аппроксимация экономической активности на территории посредством анализа космоснимков с использованием машинного обучения [62,102].

Кроме того, этот инструментарий получил распространение в урбанистике, экономике и социологии города [132,134]: благодаря продвинутым базам данных в ведущих мировых городах (в том числе Москве) на уровне мегаполиса проводятся исследования различных пространственно-временных явлений (например, активности в социальных сетях для выявления ключевых туристических объектов).

В диссертационном исследовании происходит апробация применения инструментов машинного обучения для решения научно-прикладной задачи оценки территории для размещения высокотехнологичной промышленности в России на мезоуровне.

На первом этапе решении поставленной задачи происходила обработка информационной базы исследования для подготовки к построению модели. Первичной единицей модели являются действующие высокотехнологичные предприятия, по ним были рассчитаны зависимая и независимые переменные. Так как работа велась с пространственными данными, обработка имела свою специфику и подразумевала применение следующих техник:

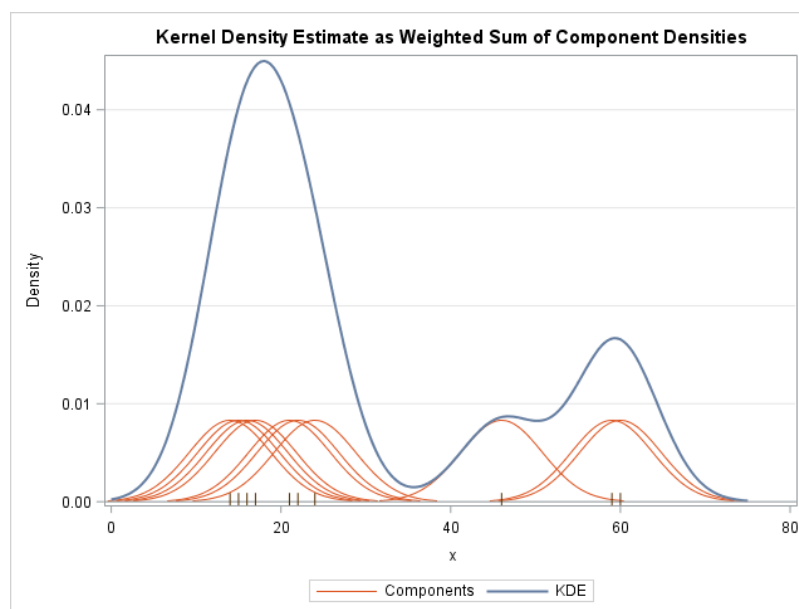


Рисунок 16 – Визуализация принципа работы ядерной плотности

Источник: How to visualize a kernel density estimate [159]

- 1) использование простой ядерной плотности. Для больших массивов точечных объектов построена взвешенная ядерная плотность. Этот метод помогает распределить эффект влияния объекта по заданной площади для построения сглаженной пространственной картины. Оценка ядерной плотности входит в число рекомендуемых инструментов для анализа

точечных паттернов для пространственного анализа [38], позволяет отойти от проблемы MAUP;

- 2) применение взвешенной ядерной плотности. Для массивов с возможностью оценки «силы» их влияния на окружающие объекты, ядерную плотность можно взвешивать. В таком случае вклад каждой точки в плотность неоднороден, чем выше значение весового признака, тем сильнее влияние каждой отдельной точки;

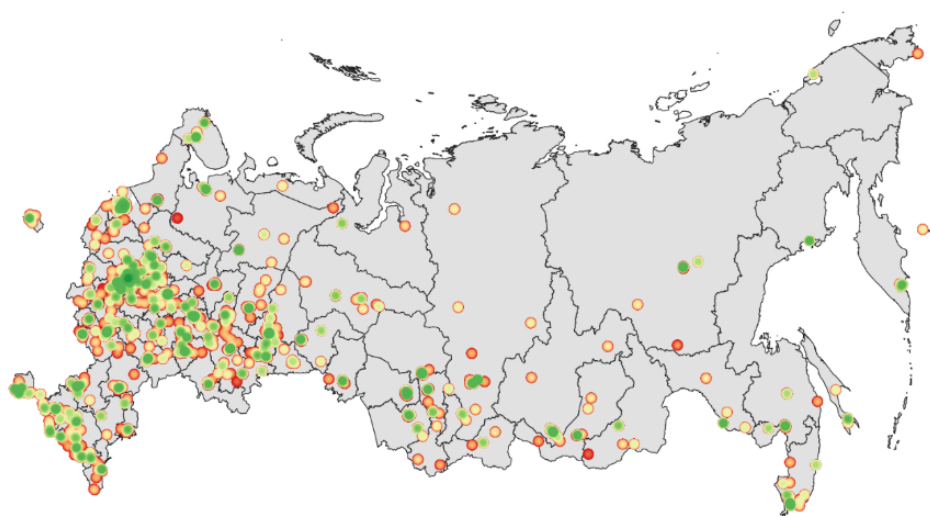


Рисунок 17 – Результат интерпретации данных о НИР через ядерную плотность, взвешенную по объему исполнения заказов

Источник: составлено автором на основе данных ЕГИСУ НИОКТР.

- 3) расчет расстояния до ближайшего объекта. Для простейших объектов (отдельных точек и полигонов) рассчитывался длина линии, проецируемой от предприятия до ближайшего объекта (или его границ) [142]. Расчет доступен с помощью базового инструмента QGIS – «Расстояние до ближайшего хаба»;
- 4) семплирование растровых данных. Для получения данных из растров (геопривязанных изображений, в каждом пикселе которых закодирована

информация), был использован специализированный инструмент привязки из набора SAGA GIS (входит в стандартный пакет QGIS);



Рисунок 18 – Визуальная интерпретация работы инструмента «Расстояние до ближайшего хаба» по отношению к морским портам

Источник: составлено автором.

- 5) нахождение внутри объекта. В случае привязки площадных данных, использовалась простая пространственная привязка «один к многим» (например, привязка численности населения по сетке муниципального деления).



Рисунок 19 – Визуализация климатических растровых данных с сетью высокотехнологичных предприятий

Источник: составлено автором.

В качестве зависимой переменной была сконструирована ядерная плотность высокотехнологичных предприятий, взвешенная по их среднегодовой выручке за 2019–2022 гг. радиусом в 20 км. По замыслу исследования этот показатель является прокси-индикатором потенциала территории для размещения ВТП. Зона в 20 километров выбрана как ориентировочное расстояние, которое можно преодолеть в городе за один час. Независимые переменные для модели представлены в Таблице 5.

Таблица 5 – Перечень факторов размещения, анализируемых в моделях машинного обучения

№ п/п	Фактор размещения	Первичный источник	Название переменной
1	Вид деятельности	СПАРК-Интерфакс	<i>okved</i> (дамми-переменная)
2	Численность населения	БДПМО	<i>population</i>
3	Средняя заработная плата в крупных и средних предприятиях	БДПМО	<i>wage</i>

4	Прибрежное положение муниципалитета нахождения	Подготовлено на основе данных OSM	<i>coast</i> (дамми-переменная)
5	Приграничное положение муниципалитета нахождения	Подготовлено на основе данных OSM	<i>border</i> (дамми-переменная)
6	Средняя температура июня	WorldClim	<i>clim_jun</i>
7	Средняя температура декабря	WorldClim	<i>clim_dec</i>
8	Расстояние до региональной столицы	Подготовлено на основе данных OSM	<i>distance_reg</i>
9	Расстояние до городов-миллионников	Подготовлено на основе данных OSM и БДМПО	<i>distance_mil</i>
10	Расстояние до основных ж/д путей	Подготовлено на основе данных OSM	<i>distance_rail</i>
11	Расстояние до основных автодорог	Подготовлено на основе данных OSM	<i>distance_road</i>
12	Расстояние до аэропортов	Подготовлено на основе открытого реестра	<i>distance_aero</i>
13	Расстояние до речных портов	Подготовлено на основе перечня речных портов	<i>distance_river</i>
14	Расстояние до морских портов	Подготовлено на основе реестра морских портов	<i>distance_sea</i>
15	Расстояние до кластеров	Подготовлено на основе данных ГИСП	<i>distance_cluster</i>
16	Расстояние до промышленных парков	Подготовлено на основе данных ГИСП	<i>distance_indp</i>
17	Расстояние до технопарков	Подготовлено на основе данных ГИСП	<i>distance_tpark</i>
18	Расстояние до специальных экономических зон (ОЭЗ, ТОСЭР, ТОР)	Подготовлено на основе открытых реестров	<i>distance_sez</i>
19	Расстояние до школ из ТОП-200	Подготовлено на основе данных ГИСП	<i>distance_school</i>

	рейтинга RAEX	основе открытых рейтингов	
20	Ядерная плотность по численности выпускников гуманитарных направлений	Подготовлено на основе данных мониторинга высшего образования ВШЭ	<i>hum_kern</i>
21	Ядерная плотность по численности выпускников естественно-научных направлений	Подготовлено на основе данных мониторинга высшего образования ВШЭ	<i>nat_kern</i>
22	Ядерная плотность по численности выпускников естественно-научных направлений	Подготовлено на основе данных мониторинга высшего образования ВШЭ	<i>tech_kern</i>
23	Ядерная плотность размещения медицинских организаций	Подготовлено на основе данных Росздравнадзора	<i>med_kern</i>
24	Ядерная плотность размещения банкоматов	Подготовлено на основе открытых реестров	<i>atm_kern</i>
25	Ядерная плотность размещения банковских отделений	Подготовлено на основе открытых реестров	<i>banks_kern</i>
26	Ядерная плотность по объему закупок НИР	Подготовлено на основе данных ЕГИСУ НИОКТР	<i>nir_buyers_kern</i>
27	Ядерная плотность по объему исполнения НИР	Подготовлено на основе данных ЕГИСУ НИОКТР	<i>nir_cond_kern</i>
28	Ядерная плотность размещения организаций среднего профессионального образования	Подготовлено на основе реестра Рособрнадзора	<i>spo_kernel</i>

Источник: составлено автором.

Дальнейший алгоритм формирования, валидации и тестирования модели в общих чертах производился следующим образом:

- 1) подгрузка необходимых библиотек (в том числе *sklearn*, *catboost*, *lightgbm*, *shap*);
- 2) загрузка и предобработка данных;
- 3) анализ данных (подготовка графиков распределений, корреляционных матриц);
- 4) подготовка тренировочной и тестовой выборок;
- 5) подбор моделей с настройкой гиперпараметров с помощью кросс-валидации;
- 6) оценка важности показателей модели с применением *shap-values*;
- 7) дополнительная итерация подбора на основе анализа *shap-values* для оптимизации вычислений;
- 8) применение модели к тестовой выборке, визуализация результатов моделирования в ГИС;
- 9) визуальная оценка качества модели, описание «хот-спотов» переоценки или недооценки модели в сравнении с актуальными данными.

После подготовки итоговой базы данных (30 910 наблюдений, 29 показателей) был проведен пробный запуск модели дерева решений с подбором гиперпараметров, который показал чрезмерно высокие показатели метрик на всех выборках данных, и высокий уровень гетероскедастичности предсказаний.

Оценка проблемы выявила своеобразную пространственную утечку данных связанную с гиперконцентрацией практически всех показателей (в том числе зависимой переменной) в Москве и Петербурге. В связи с этим дальнейший анализ проводился без наблюдений в этих городах (таким образом, мы априорно предполагаем, что они подходят для размещения ВТП), по 22 206 наблюдениям. Дальнейший ход моделирования представлен далее в тексте параграфа, а также в виде кода с комментариями в Приложении 1.

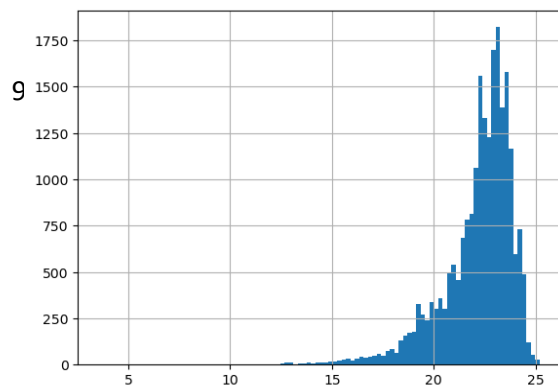
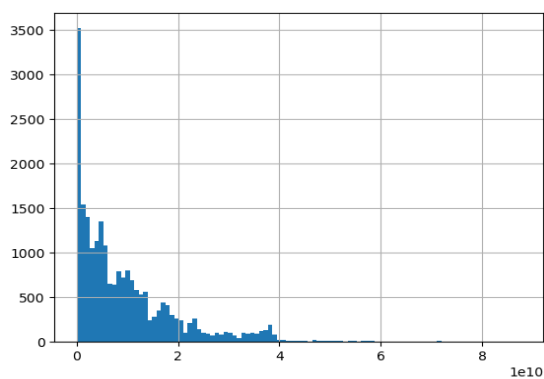


Рисунок 20 – Гистограмма распределения зависимой переменной без данных Москвы и Санкт-Петербурга (слева – без обработки, справа – с нормализацией по натуральному логарифму)

Источник: составлено автором

В ходе предобработки данных были построены гистограммы распределения для зависимой и независимых переменных. Для упрощения работы алгоритмов машинного обучения зависимая переменная, а также переменные закупок и исполнения НИР были взяты по натуральному логарифму их значений (с созданием новых переменных *inc_kernel_ln*, *nir_buyers_ln* и *nir_cond_ln* соответственно).

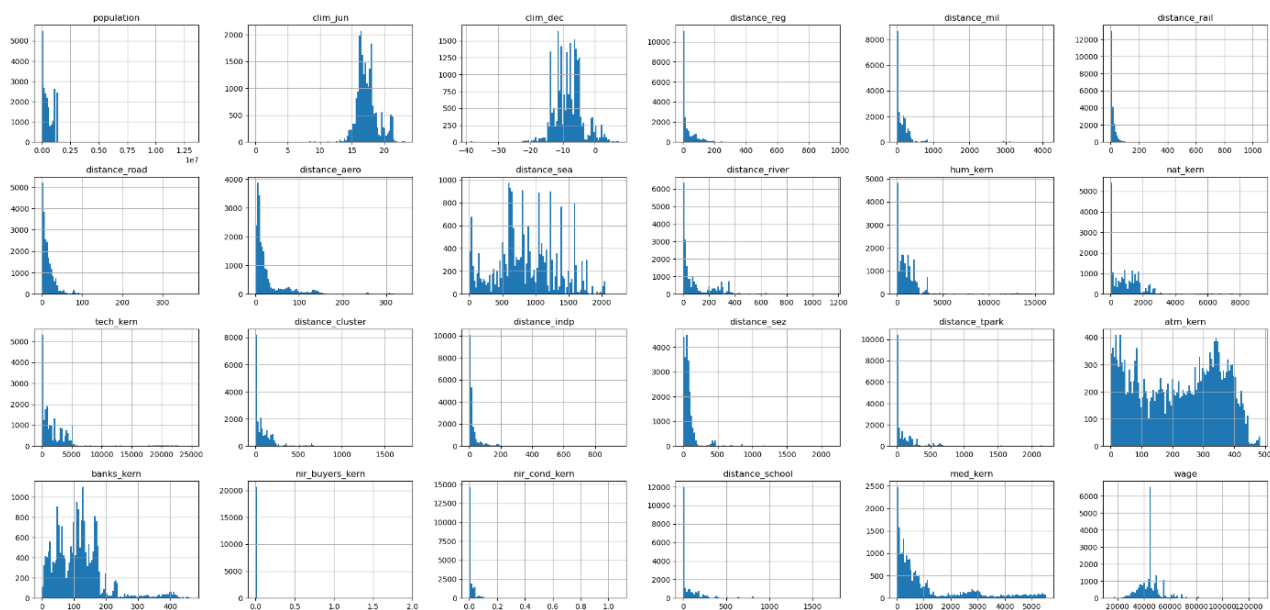


Рисунок 21 – Гистограммы распределения зависимой и независимых переменных

Источник: составлено автором.

Корреляционный анализ показал высокий уровень мультиколлинеарности в данных, в связи с чем ряд переменных подвергся дополнительной обработке. В

первую очередь переменные численности выпускников вузов по разным специальностям были объединены в общую переменную путем суммирования (полученная синтетическая переменная – *univ_kern*). Высокая корреляция показателей объясняется «классическим» характером высшего образования в России. Аналогичная ситуация с заказчиками и исполнителями НИР, по их сумме подготовлен показатель *nir_kern_ln* – для системы НИР частым случаем является участие крупных исследовательских организации как в роли заказчика, так и в роли исполнителя исследований.

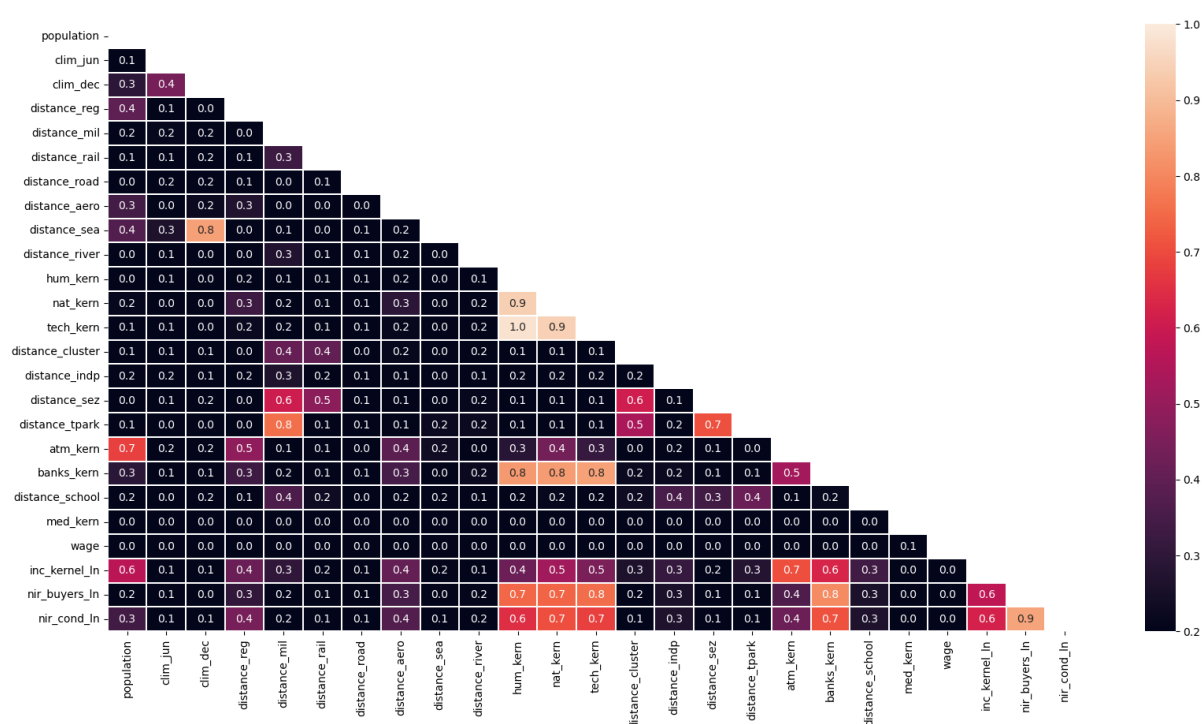


Рисунок 22 – Матрица корреляции Пирсона для переменных

Источник: составлено автором.

На следующем этапе общая база была разделена на тренировочный и тестовый наборы в пропорции 3 к 1. Тренировочный набор был использован в применении к пяти моделям – базовой дамми-модели (предсказание медианным значением), модели дерева решений, модели случайного леса, а также двум моделям, использующим технику градиентного бустинга – LGBM и Catboost). В ходе подбора модели была использована кросс-валидация для предотвращения

переобучению модели, оптимизация модели происходила с помощью коэффициента детерминации (R^2), вспомогательной являлась среднеквадратичная ошибка (RMSE). В ходе кросс-валидации происходил подбор глобальных (общих для всех моделей) и локальных (специфичных) гиперпараметров. Под глобальными гиперпараметрами подразумевается применение метода главных компонент для снижения размерности ряда независимых переменных и нормирования.

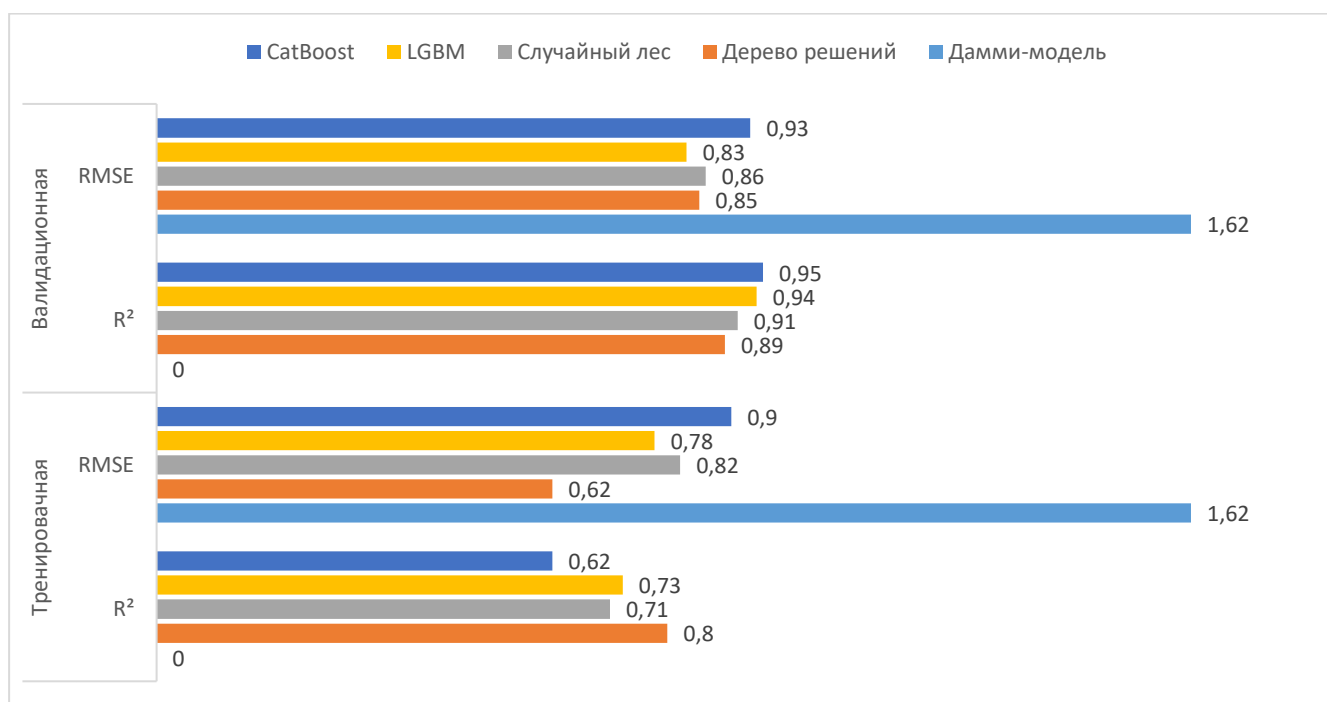


Рисунок 23 – Результаты тренировки и кросс-валидации

Источник: составлено автором.

Дамми-модель имеет низкую предсказательную способность. На тренировочной выборке лучший результат показывает более простая модель дерева решений, однако на валидации лучшими являются модели с градиентным бустингом – по метрике R^2 лучше всего работает CatBoost.

На основе SHAP-value – метрики значимости отдельных факторов модели, основанной на подходах теории игр, была произведена оценка факторов, влияющих на подготовленную ранее модель. Наибольшую значимость имеет

плотность размещения банкоматов – прокси-индикатор экономической активности на территории и проникновения банковских услуг. В топ факторов также входят переменная ядерной плотности по численности выпускников, ядерная плотность размещения банков, численность населения и близость к технопаркам.

После была сформирована еще одна CatBoost модель из набора факторов с наибольшим весом на основе анализа SHAP и лучшей интерпретируемостью для попытки оптимизации качества и скорости работы модели – *atm_kern*, *univ_kern*, *population*, *distance_tpark*, *distance_school*. В ходе кросс-валидации модель достигла оценки R^2 в 0,92 и RMSE в 0,87.

В итоге для тестирования была выбрана модель на полной выборке переменных: по результатам проверки на тестовом наборе данных она достигла R^2 в 0,92 и RMSE в 0,2.

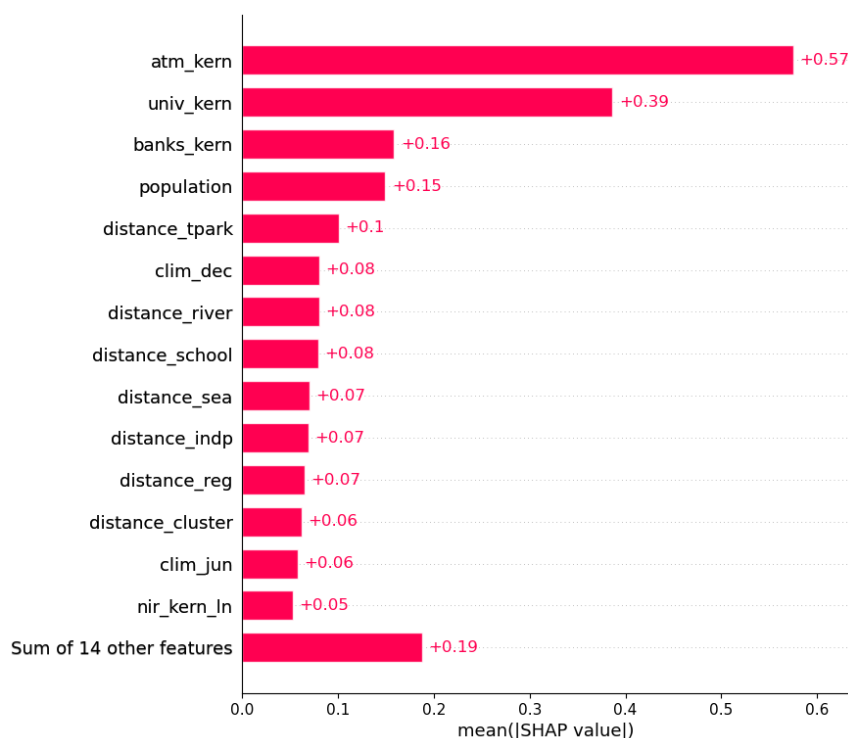


Рисунок 24 – График значимости переменных для работы модели с лучшими результатами

Источник: составлено автором.

В качестве экспертной валидации результаты моделирования по полному набору переменных были визуализированы с помощью ГИС (Рисунок 24). В целом модель хорошо справляется с предсказаниями по малым и средним городам, но завышает значения для региональных столиц (Воронежа, Самары, Кирова, Перми, Иркутска и т.д.), что можно интерпретировать в нескольких суждениях. Первое – региональные столицы имеют потенциал к развитию высокотехнологичного бизнеса и дальнейшего наращивания активности ВТП. Второе – в региональных столицах действуют высокотехнологичные предприятия, которые не входят в периметр анализа (т.е. предприятия ОПК) либо не отразились в базе, будучи структурами в рамках крупных холдингов (ОАК, Ростеха, Росатома и т.д.).

Интересной закономерностью являются заниженные предсказания в центрах химической промышленности – Дзержинске, Тольятти, Новомосковске, Набережных Челнах, Альметьевске, Уфе, Стерлитамаке и др. Несмотря на то, что модель учитывала отраслевую принадлежность бизнеса, исходя из качественного анализа результатов можно предположить, что сформированный набор факторов не позволил отразить потенциал размещения средне- и крупнотоннажной химии, в том числе из-за высокой важности «реликтовых» площадок, экологического и общественно-политического факторов для этого направления.

Можно также предположить недооценку влияния особых экономических зон моделью (включая ТОРы и ТОСЭРы): помимо уже упомянутых территорий, заниженные значения в сравнении с актуальными имеют районы Ульяновска, Дубны, Кондрово и др. Seriously недооценен потенциал Екатеринбурга и Волгограда, что может быть связано с ошибками в обработке входных данных или их качестве. Итоговая модель позволяет решать сразу несколько задач. Во-первых, модель позволяет оценить наиболее благоприятные локации для предприятий разных видов деятельности в рамках группы высокотехнологичных производств, т.е. применима для управления региональной сетью бизнес-структур. Во-вторых, модель применима для оценки изменения среды: появление новых ОЭЗ, университетов, школ влияет на модель и позволяет оценить рост

потенциала территории для размещения, т.е. может быть использована в вопросах государственного территориального планирования. «Побочным» итогом моделирования является выявление территорий по тем или иным причинам, не развивающим ВТП, несмотря на имеющиеся предпосылки.

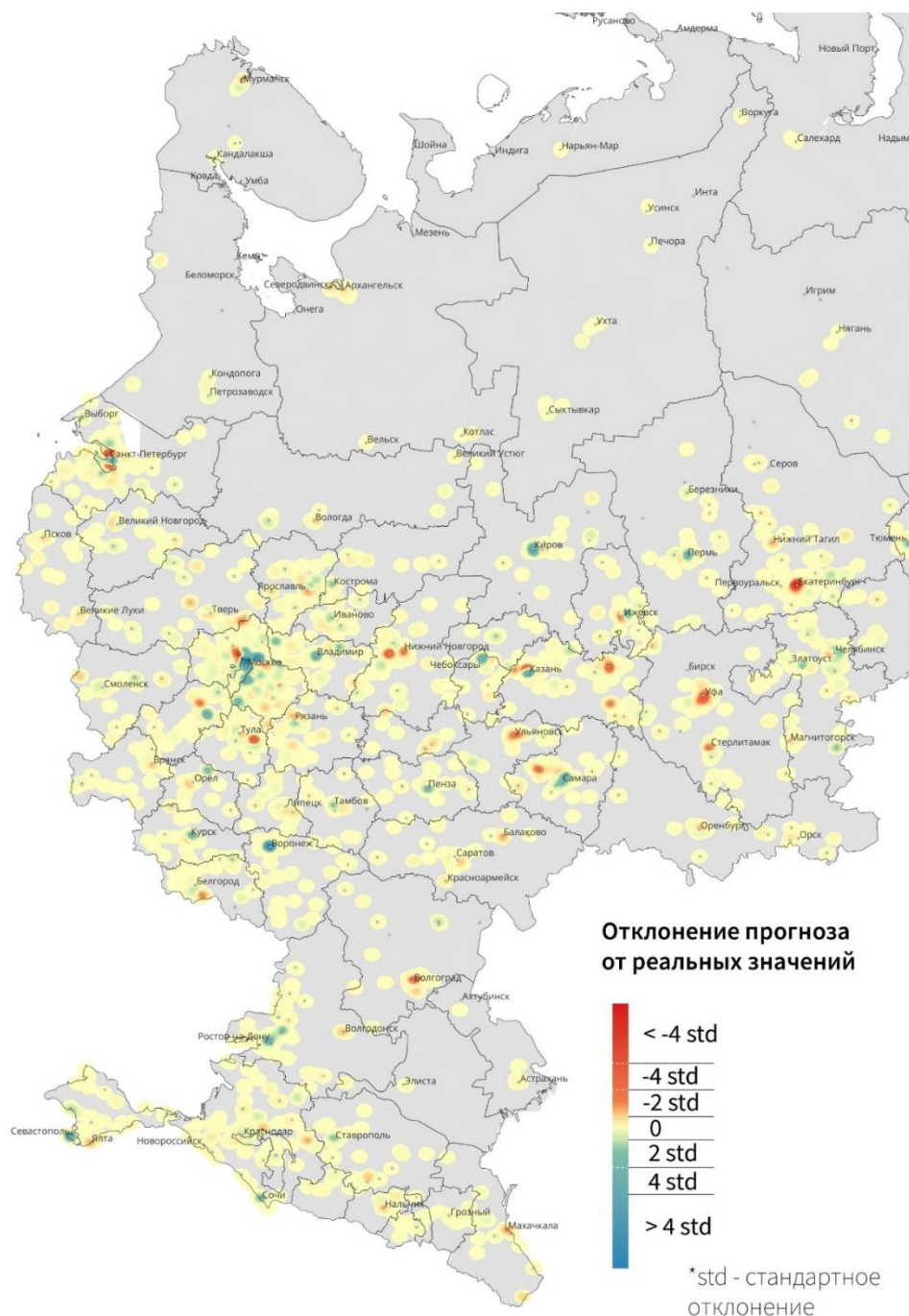


Рисунок 25 – Разница между предсказанными и актуальными значениями в формате тепловой карты

Источник: составлено автором.

В рамках параграфа представлена апробация использования машинного обучения для оценки потенциала территории к размещению ВТП. Сравнительно высокая предсказательная способность подготовленной модели и адекватность ее предсказаний позволяют подтвердить возможность применения моделей машинного обучения для решения вопросов размещения на мезоуровне, а также для выявления ключевых факторов размещения.

3.4 Формирование оптимальной модели выбора локации

Предыдущие этапы анализа раскрыли особенности задач выбора локации ВТП. Подготовленные ранее тезисы и выводы, выявленные закономерности и проведенная апробация различных методов и инструментов позволяют сформировать общий оптимальный алгоритм поиска локации для высокотехнологичного предприятия. Несмотря на то, что алгоритм был апробирован на уровне муниципалитетов и крупных агломераций, предполагается, что возможно его использование уже на уровне конкретных локаций (микроуровне) с соответствующими изменениями к составу факторов размещения.

Критичной является роль эксперта/экспертной группы, которая решает вопросы размещения. Экспертное знание является ключевым для финального выбора локации, так как может учесть неявные или специализированные факторы размещения: институциональные и политические, технико-экономические, экологические. Без экспертов также невозможно формирование первичного набора факторов размещения. В целом на основе предложенного алгоритма возможно формирование отдельного «пути» решения задачи размещения, базирующегося исключительно на экспертных процедурах (Рисунок 26).



Рисунок 26 – Базовый алгоритм выбора локации для компаний ВТП на основе экспертного знания.

Источник: составлено автором.

При этом необходимо понимать ограничения экспертных методов и возможные ошибки, связанные с недостаточной осведомленностью экспертной группы (особенно если речь идет о совокупности отраслей в масштабах всей страны). В таком случае экспертное знание может быть дополнено и/или валидировано с помощью интерактивных систем принятия решений (Рисунок 27) с визуализированными массивами информации об отраслевых особенностях и потенциальных факторах размещения (в том числе в формате пространственных данных) на разных масштабных уровнях.



Рисунок 27 – Комбинированный алгоритм выбора локации для компаний ВТП с применением СППР.

Источник: составлено автором.

Наконец, экспертное знание может быть дополнено результатами работы алгоритмов машинного обучения (Рисунок 28). На основе подготовленного экспертами набора факторов, применимого для прогнозирования пространственной конфигурации отрасли, а также пригодного к оцифровке, формируется модель машинного обучения, которая оценивает потенциал территории для размещения бизнеса, исходя из выбранной метрики (выручка, численность занятых, количество предприятий, продаж и т.д.). Опционально модель может быть основана на информации из СППР, что потенциальнократно упрощает процесс сбора и обработки данных.

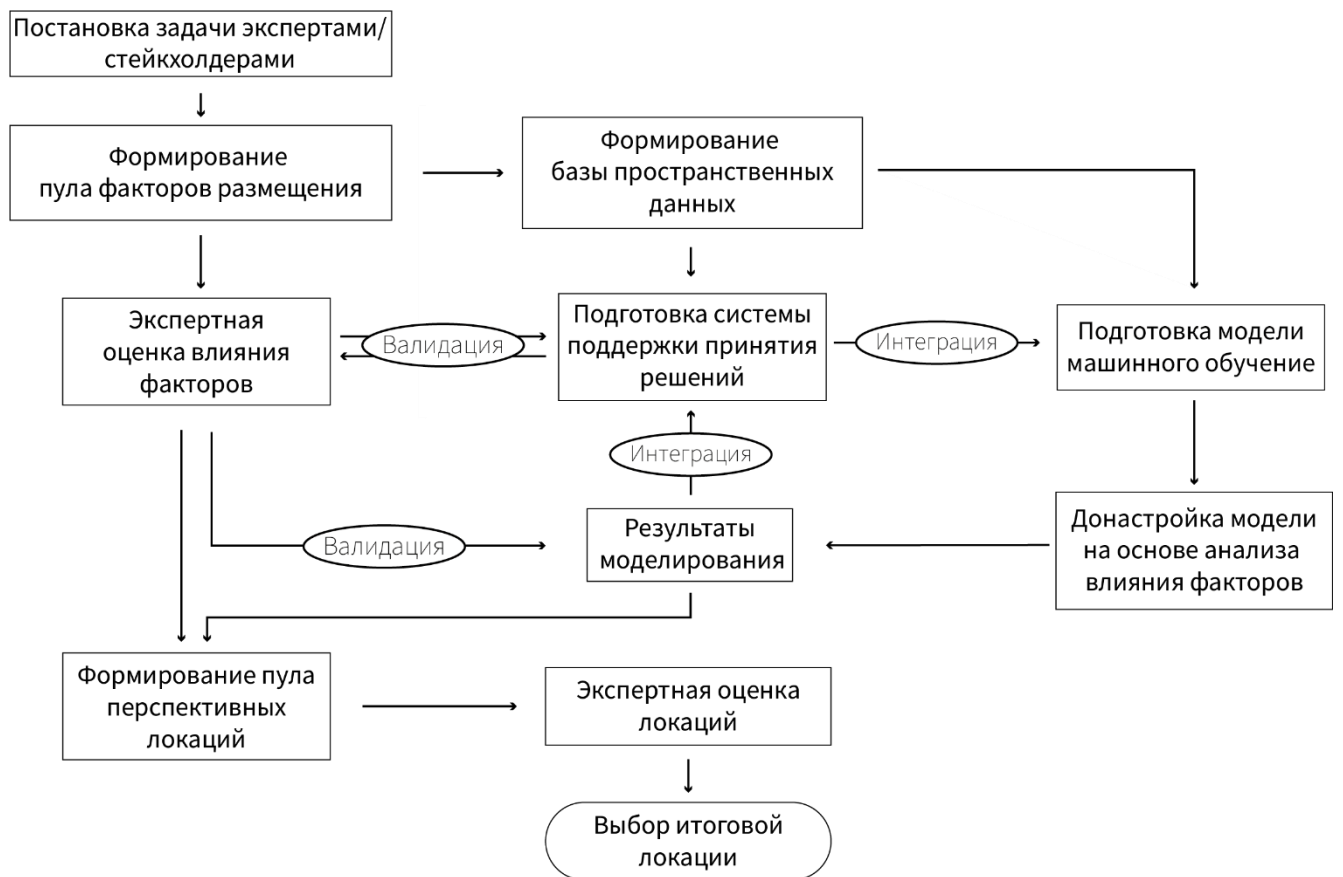


Рисунок 28 – Комплексный алгоритм выбора локации для компаний ВТП

Источник: составлено автором.

По результатам моделирования готовится перечень локаций, наиболее благоприятных для размещения предприятия – результаты моделирования также могут быть интегрированы в СППР для удобства работы экспертов. Результаты работы модели проходят валидацию, в случае адекватности предсказаний могут расширить набор локаций после экспертного выбора. Несмотря на то, что в работе рассматривалось только применение машинного обучения, эконометрическое моделирование также может быть встроено в алгоритм выбора локации.

Таким образом, в зависимости от полноты данных, компетенций исследовательской или аналитической команды, временных ресурсов на решение задач размещения можно рассматривать сразу несколько алгоритмов поиска оптимальной локации:

- 1) Базовый – с применением поэтапного анализа на основе экспертного знания. Такой алгоритм наиболее применим для решения задач размещения при заведомом понимании отраслевых особенностей и тонкостей, наиболее эффективен для решения задач на уровне поиска локации внутри одного или небольшой группы регионов;
- 2) Комбинированный – с применением СППР для валидации и поддержки экспертного знания. В рамках этого алгоритма к использованию экспертного знания добавляется использование цифровых инструментов (в частности интерактивных аналитических панелей) для уточнения предложений по оптимальной локации, позволяет обработать и визуализировать большой объем данных. В связи с этим, такой подход представляется подходящим для анализа в рамках групп регионов (например, федеральных округов);
- 3) Комплексный – с применением моделирования (машинного обучения или эконометрических моделей). Такой метод в связке с СППР позволяет снизить нагрузку на экспертный блок, выявить неявные площадки для размещения высокотехнологичных предприятий. Этот алгоритм оптимален при работе в рамках всей страны.

Выводы по итогам главы. Обширный научно-практический инструментарий для решения задач размещения производств условно делится на три направления – использование экспертных методов, эконометрических моделей и машинного обучения. Все три направления встречаются как в теоретических, так и в практических работах, дискутируются преимущества одних инструментов над другими, их недостатки.

Исходя из синтеза теоретических знаний, полученных при анализе литературы, предполагается, что наиболее эффективной является композиция из нескольких инструментальных направлений для формирования механизмов кросс-валидации и выявления неявных закономерностей. Однако в таком случае критическим становится вопрос обеспеченности данными, необходимыми для обеспечения нормальной работы количественного инструментария.

В главе представлены возможности интеграции микроданных из разных источников в VI-платформу для решения задач размещения, апробирована концепция применения дашбордов в качестве СППР для задач планирования размещения, валидации экспертных гипотез, визуализации и хранения данных для машинного обучения. Использование VI-инструментов (в том числе облачных геоинформационных систем) – перспективное направление совершенствования региональной промышленной и инвестиционной политики в части оценки потенциала территории для размещения ВТП. Так, ряд регионов (например, Калужская область) уже использует ГИС для мониторинга и продвижения территории с целью размещения новых предприятий.

Также проведена апробация методов анализа размещения на основе использования методов машинного обучения. Подход к размещению с использованием методов машинного обучения показал свою состоятельность: несмотря на обширный территориальный и отраслевой охват, а также связанную с этим неоднородность данных, сформированная в ходе исследования модель имеет высокую предсказательную способность. Географическая интерпретация остатков модели указала на ее слабые места: переоценку региональных столиц и недооценку ключевых центров химической промышленности, однако в целом в масштабах страны предсказания модели являются адекватными и отражают актуальную территориальную структуру ВТП.

На основе предшествующих элементов исследования был подготовлен оптимальный алгоритм выбора локации, исходя из спектра методов, представленных и апробированных в работе. Алгоритм применим для разных масштабных уровней, позволяет выстроить пошаговую работу для решения задачи размещения.

ЗАКЛЮЧЕНИЕ

Вопрос выбора локации предприятием является одним из ключевых для его стратегии становления и развития. Научно-практическая база по теме размещения насчитывает большое количество работ за более чем столетний период, перечень современных работ обширен и разнороден, тематика характеризуется ярко выраженной междисциплинарной природой. В российской научной среде вопросы размещения являются второстепенной тематикой, несмотря на актуальность рассмотрения через призму существующих технико-экономических проблем, в основном исследования проводятся на региональном уровне или в формате анализа отдельных примеров.

Ключевым элементом при выборе локации являются факторы размещения. Проведенные исследования позволяют сформировать отдельные группы факторов, применимых для анализа высокотехнологичного бизнеса: научная инфраструктура, человеческий капитал и качество жизни, производственная и логистическая инфраструктура, средовые факторы, взаиморазмещение бизнеса и группу прочих факторов (экология, природные риски и т.д.).

На основе историко-географического и историко-экономического анализа также была проведена оценка релевантности факторов размещения по отношению к российской действительности, изучена эволюция подходов к размещению ВТП в СССР, США, Европейском Союзе, Китае. Выявлена критическая роль особых экономических зон, технопарков и университетов в развитии высокотехнологичного бизнеса на территории.

Рассмотренный на основе научных публикаций инструментарий для выбора локаций обширен, включает интерпретацию экспертных суждений, использование эконометрического и финансово-экономического моделирования, машинного обучения. В рамках работы рассмотрены методы экспертного анализа размещения и машинного обучения для выбора локации. Для валидации экспертного знания и подготовки количественных моделей ключевым является

фактор наличия и доступности информации. Развитие концепций открытых данных, появление агрегаторов микроданных позволяют формировать сложные базы, применимые для анализа размещения с точностью, недоступной исследователям ранее.

Рассмотрение ареалов размещения высокотехнологичного бизнеса на основе анализа пространственных данных позволяет выделить отдельные факторы размещения, актуальные для высокотехнологичного бизнеса на текущий момент. Вместе с тем экспертная оценка размещения указывает на объективную невозможность квантификации всех факторов размещения, что делает экспертную валидацию результатов любого количественного исследования необходимостью.

В первую очередь очевидно влияние агломерационного фактора и сопряженных с ним средовых факторов. При этом внутри агломераций на микроуровне есть своя территориальная диверсификация по уровню привлекательности для размещения ВТП.

Оценка размещения крупнейших предприятий в каждом из выделенных ареалов подтверждает влияние «реликтовых» промышленных площадок и производств, сохранившихся после распада СССР. Унаследованные крупные промышленные производства могут играть разную роль в функционировании высокотехнологичных производств, являясь ядрами концентрации для технологических «спин-оффов».

Размещение высокотехнологичных производств вне основных ареалов размещения характеризуется наличием доминирующей отрасли-заказчика продукции ВТП: например, нефтегазовой и металлургической отраслей, сельского хозяйства. В этом случае малые и средние предприятия вклиниваются в сложную цепочку поставщиков крупных промышленных предприятий.

Центрами развития новых высокотехнологичных производств выступают особые экономические зоны, технологические и индустриальные парки. Показательные примеры в виде ОЭЗ Дубна и Алабуга подтверждают эффективность концепции развития ВТП через преференциальные зоны,

использующейся в ключевых странах-лидерах по развитию высокотехнологичных производств.

В качестве одного из путей оптимизации при работе с задачами размещения в работе рассматривается построение системы поддержки принятия решений на основе открытых данных (включая пространственные индикаторы). В рамках исследования было апробировано формирование СППР для визуализации научной инфраструктуры, инфраструктуры водного транспорта и рынка фармацевтической продукции. Результаты апробации позволяют судить о возможности реализации комплексных решений по агрегированию данных из разных источников для поддержки аналитика-исследователя при оценке потенциала территории для размещения бизнеса. СППР также может являться информационной базой для подготовки решений по количественной оценке потенциала размещения, в том числе на основе методов машинного обучения.

Применение машинного обучения для решения пространственных задач нашло распространение в работах ученых естественно-научных направлений – физико-географов, биологов, почвоведов, геологов. В социальных науках, а также при решении задач размещения более распространенным остается применение эконометрических моделей.

В рамках работы проведена апробация применения машинного обучения для оценки потенциала территории для размещения высокотехнологичного бизнеса. На основе оцифрованных факторов размещения была подготовлена модель, предсказывающая потенциал территории для размещения ВТП с учетом отраслевой привязки бизнеса. Аппроксимацией потенциала является территориально взвешенная методом ядерной плотности среднегодовая выручка действующих высокотехнологичных предприятий.

Высокие значения коэффициента детерминации на тестовой выборке позволяют утверждать, что машинное обучение применимо для оценки территории для размещения бизнеса, однако со своими ограничениями. Картографическая визуализация остатков модели показала наличие закономерностей в пере- и недооценке моделью отдельных территорий.

Например, недооценка была выявлена для центров химической промышленности: предполагается, что это указывает на особые требования крупных химических предприятий к размещению (в том числе наличие «реликтовых» производственных площадей). В отличие от качественного анализа модель также недооценивала потенциал особых экономических зон и территорий опережающего развития.

По итогам работы был подготовлен алгоритм принятия решений, использующий последовательное применение апробированных инструментов. Сформировано несколько модульных треков решения вопросов размещения:

- базирующийся только на знании эксперта или экспертной группы с использованием или без использования методов обработки экспертного мнения (системы весов, аналитических моделей);
- с применением интерактивных панелей (СППР) для валидации и поддержки экспертного знания;
- с применением машинного обучения для расширения перечня потенциальных локаций на основе выявления неявных закономерностей с дальнейшей экспертной валидацией, а также опциональной интеграцией с СППР для дальнейшего улучшения модели и визуализации результатов ее работы.

Таким образом, в работе произведены апробация и синтез использования библиографического, регионального и пространственного анализа, анализа кейсов и экспертного знания, а также применения микро- и пространственных данных, интерактивных аналитических панелей и машинного обучения для формирования методологии решения задач размещения высокотехнологичной промышленности.

Работа раскрывает все положения, выносимые на защиту и подробно останавливается на каждом из них. То есть, по результатам диссертационного исследования:

1. Выявлены факторы размещения, наиболее применимые для анализа территориальной структуры высокотехнологичных предприятий;

2. Выявлены ключевые ареалы размещения ВТП на основе визуализации пространственных данных. Выделены архетипичные стратегии размещения предприятий российского высокотехнологического сектора;
3. Разработан и апробирован подход к интеграции микроданных из разных источников на базе платформы бизнес-аналитики для решения задач размещения высокотехнологических предприятий в России;
4. Разработана и апробирована модель машинного обучения для решения задач размещения высокотехнологических предприятий на уровне муниципалитетов и крупных агломераций;
5. Сформирован алгоритм принятия решений о размещении предприятий, использующий последовательное применение апробированных методологических инструментов. Сформировано несколько модульных подходов к решению вопросов размещения.

Исследование сталкивается с объективными ограничениями информационного периметра (в частности – условность включения предприятий в перечень высокотехнологических на основе ОКВЭД2), а также полноты, достаточности и достоверности данных. По результатам количественной части исследования можно предположить, что при включении предприятий химического комплекса в периметр анализа размещения высокотехнологических предприятий, стоит вывести крупнейшие химические комплексы как объекты, имеющие предпосылки для размещения кардинально иные в сравнении с общей совокупностью ВТП.

Исследование имеет высокую научную новизну для российского научного поля, применяя междисциплинарный подход с использованием методик и инструментов региональной экономики, социально-экономической географии, менеджмента. Результаты исследования могут быть востребованы как со стороны государства для оценки потенциала и поиска решений для развития ВТП в регионах России, так и со стороны бизнеса для формирования набора локаций, привлекательных для размещения их предприятий.

Дальнейшими шагами для исследования в рамках тематики может быть анализ не статичных, а пространственно-временных территориальных конструкций с применением современных подходов бизнес-демографии, что потребует формирования более сложных баз панельных данных.

В рамках научной или практической деятельности также возможна апробация всего контура предложенного алгоритма выбора локации, включая интеграцию машинного обучения с СППР, и подготовки системы парсеров, оперативно обновляющих информационную базу исследования, а также ее обогащение с использованием информации проектируемой Национальной системы пространственных данных. Наконец, развитие исследования возможно в проведении мультимасштабного анализа факторов размещения от мезо- к микроуровню (уровню конкретных локаций) с формированием и оценкой иерархии факторов размещения, в том числе с применением методов форсайт-исследований.

СПИСОК ЛИТЕРАТУРЫ

- 1) Анимица Е. Г. Импортозамещение в промышленном производстве региона: концептуально-теоретические и прикладные аспекты / Е. Г. Анимица, П. Е. Анимица, А. А. Глумов // Экономика региона. — 2015. — Том 3. — С. 160–172.
- 2) Бабурин, В. Л. Влияние эффекта колеи на эволюцию промышленных ареалов России / В. Л. Бабурин // Региональные исследования. — 2020. — № 3. — С. 26–39.
- 3) Бабурин, В. Л. Московский столичный регион на рубеже веков: новейшая история и пути развития / В. Л. Бабурин, В. Р. Битюкова, М. А. Казьмин, А. Г. Махрова. — Смоленск : Ойкумена. 2003.
- 4) Бабурин, В. Л. Проблемы формирования территориально-отраслевой структуры промышленности Москвы / В. Л. Бабурин, А. Г. Махрова // Географические исследования в Московском регионе. Москва 1987. — С. 39–44.
- 5) Бабурин, В. Л. Регионы-новаторы и инновационная периферия России. Исследование диффузии инноваций на примере ИКТ-продуктов / В. Л. Бабурин, С. П. Земцов // Региональные исследования. — 2014. — № 3(45). — С. 27–37.
- 6) Бабурин, В. Л., Земцов, С.П. География инновационных процессов в России // Вестник Московского университета. Серия 5. География. — 2013. — №5. — С. 25–32.
- 7) Бакланов, П. Я. Территориальная организация хозяйства и региональное развитие / П. Я. Бакланов // Вестник Арго. — 2012. — № 1. — С. 32–49.
- 8) Горлов, В. Н. Москва в контексте процесса деиндустриализации / В. Н. Горлов // Глобальный город: теория и реальность. — М., 2007. — С. 198–208.

- 9) Гранберг, А.Г. Основы региональной экономики. – М.: Государственный университет – Высшая школа экономики – 2001.
- 10) Гурков, И. Б. Размещение российских предприятий зарубежных корпораций, открытых в 2012–2018 гг. / И. Б. Гурков // Пространственная экономика. — 2019. — № 2. — С. 17–36.
- 11) Дружинин А. Г., Колосов В. А., Шувалов В. Е. Пространство современной России: возможности и барьеры развития (размышления географов-обществоведов) / А. Г. Дружинин, В. А. Колосов, В. Е. Шувалов. — М. : Вузовская книга, 2012. — 336 с.
- 12) Еферин, Я. Ю., Куценко, Е. С. Адаптация концепции умной специализации для развития регионов России // Вопросы государственного и муниципального управления. — 2021. — №3. — С. 75–110.
- 13) Жаркова, Е. С. Экономические теории размещения производства: от штандорта к кластерам / Е. С. Жаркова // Вестник Санкт-Петербургского университета. Экономика. — 2011. — № 1. — С. 145–150.
- 14) Земцов, С. П. Потенциальные высокотехнологичные кластеры в российских регионах: от текущей политики к новым точкам роста / С. П. Земцов, В. А. Барина, А. А. Панкратов, Е. С. Куценко // Форсайт. — 2016. — Т. 10, № 3. — С. 34–52.
- 15) Клочков, В. В. Пространственные и поселенческие аспекты инновационного развития наукоемкой промышленности в России / В. В. Клочков, Е. Ю. Байбакова // Национальные интересы: приоритеты и безопасность. — 2011. — № 42.
- 16) Колосовский, Н. Н. К вопросу об экономическом районировании / Н. Н. Колосовский // Пространственная экономика. — 2009. — № 1. — С. 102–123.
- 17) Кузнецова, О. В. Факторы размещения фармацевтических предприятий на примере крупнейших фармкомпаний России / О. В. Кузнецова, Р. О. Бобровский // Известия Российской академии наук. Серия географическая. — 2022. — Т. 86, № 2. — С. 168–178.

- 18) Литовский, В. В. К проблеме новой индустриализации и перехода к очередным технологическим укладам на Урале и сопряженных территориях. Часть 2. Советские практики // История и современное мировоззрение. — 2022. — №3. — С. 26–40
- 19) Максименко, Д. Д. Актуальные факторы размещения высокотехнологичных производств в России / Д. Д. Максименко // Региональные исследования. — 2023. — № 4. — С. 29–40.
- 20) Манаков, А. Г. «Новая экономическая география» и оценка ее применимости в России / А. Г. Манаков // Вестник Псковского государственного университета. Серия: Естественные и физико-математические науки. — 2012. — № 1.
- 21) Михайлов, А. С. Научно-инновационная динамика развития Северо-Запада России в условиях геополитических перемен / А. С. Михайлов, Д. Д. Максименко, М. Р. Максименко, М. М. Филатов // Балтийский регион. — 2023. — Т. 15, № 4. — С. 79–103.
- 22) Михайлов А. С., Максименко Д. Д., Максименко М. Р. Пространственные и структурные закономерности в распределении научно-технологической, инновационной и производственной деятельности в России // Балтийский регион. 2024. – № 2. – С. 41–62.
- 23) Саушкин, Ю. Г. Основные экономические районы СССР / Ю. Г. Саушкин, Т. М. Калашникова // Вопросы географии. — 1959. — № 47. — С. 42–73.
- 24) Хрущев, А. Т. География промышленности СССР / А. Т. Хрущев. — М. : Мысль, 1969.
- 25) Чернышова, Н. А. Современные проблемы реорганизации производственных зон Москвы / Н. А. Чернышова // Вестник Московского университета. Серия 5: География. — 2010. — № 1. — С. 72–77.
- 26) Шарыгин, М. Д. Эволюция учения о территориальных общественных системах / М. Д. Шарыгин // Географический вестник. — 2006. — № 1. — С. 4–13.

- 27) Юсупова, А. Т., Халимова, С. Р. Высокотехнологичный бизнес в регионах России: роль в экономике, дифференциация и основные детерминанты развития // Вестник Санкт-Петербургского университета. Менеджмент. 2020. №1. — С. 67–95.
- 28) Abramovsky, L. University research and the location of business R&D / L. Abramovsky, R. Harrison, H. Simpson // *The Economic Journal*. — 2007. — Vol. 117, No. 519. — P. 114–141.
- 29) Allio, M. K. Strategic dashboards: designing and deploying them to improve implementation / M. K. Allio // *Strategy & Leadership*. — 2012. — Vol. 40, No. 5. — P. 24–31.
- 30) Andrienko, G. Geovisual analytics for spatial decision support: Setting the research agenda / G. Andrienko, N. Andrienko, P. Jankowski, D. Keim, M. J. Kraak, A. MacEachren, S. Wrobel // *International journal of geographical information science*. — 2007. — Vol. 21, No. 8. — P. 839–857.
- 31) Anselin, L. Local geographic spillovers between university research and high technology innovations / L. Anselin, A. Varga, Z. Acs // *Journal of urban economics*. — 1997. — Vol. 42, No. 3. — P. 422–448.
- 32) Appold, S. Location patterns of US industrial research: mimetic isomorphism and the emergence of geographic charisma / S. Appold // *Regional studies*. — 2005. — Vol. 39, No. 1. — P. 17–39.
- 33) Arauzo-Carod, J. M. Empirical studies in industrial location: an assessment of their methods and results / J. M. Arauzo-Carod, D. Liviano-Solis, M. Manjón-Antolín // *Journal of Regional Science*. — 2010. — Vol. 50, No. 3. — P. 685–711.
- 34) Arentze, T. A. The integration of expert knowledge in decision support systems for facility location planning / T. A. Arentze, A. W. Borgers, H. J. Timmermans // *Computers, Environment and Urban Systems*. — 1995. — Vol. 19, No. 4. — P. 227–247.
- 35) Armstrong, M. P. A knowledge-based approach for supporting locational decisionmaking / M. P. Armstrong, S. De, P. J. Densham, P. Lolonis, G. Rushton,

- V. K. Tewari // *Environment and Planning B: Planning and Design*. — 1990. — Vol. 17, No. 3. — P. 341–364.
- 36) Athreye, S. S. Role of Transnational Corporations in the Evolution of a High-Tech Industry: The Case of India's Software Industry — A Comment / S. S. Athreye // *World Development*. — 2004. — Vol. 32, No. 3. — P. 555–560.
- 37) Audretsch, D. B. University spillovers and new firm location / D. B. Audretsch, E. E. Lehmann, S. Warning // *In Universities and the Entrepreneurial Ecosystem*. — Edward Elgar Publishing, 2017.
- 38) Baddeley, A. Spatial point patterns: methodology and applications with R / A. Baddeley, E. Rubak, R. Turner. — CRC press, 2015.
- 39) Badri, M. A. Combining the analytic hierarchy process and goal programming for global facility location-allocation problem / M. A. Badri // *International Journal of Production Economics*. — 1999. — Vol. 62, No. 3. — P. 237–248.
- 40) Badri, M. Dimensions of industrial location factors: Review and exploration / M. Badri // *Journal of Business and Public Affairs*. — 2007. — Vol. 1. — P. 1–27.
- 41) Balbontin, C. Firm-specific and location-specific drivers of business location and relocation decisions / C. Balbontin, D. A. Hensher // *Transport Reviews*. — 2019. — Vol. 39, No. 5. — P. 569–588.
- 42) Bivand, R. A review of spatial statistical techniques for location studies / R. Bivand // *CEPR symposium on new issues in trade and location*. — Lund, 1998. — P. 1–21.
- 43) Blair, J. P. Major factors in industrial location: A review / J. P. Blair, R. Premus // *Economic development quarterly*. — 1987. — Vol. 1, No. 1. — P. 72–85.
- 44) Boschma, R. A. Evolutionary economics and industry location / R. A. Boschma, K. Frenken // *Review for Regional Research*. — 2003. — Vol. 23. — P. 183–200.
- 45) Brunnermeier, S. B. Examining the evidence on environmental regulations and industry location / S. B. Brunnermeier, A. Levinson // *The Journal of Environment & Development*. — 2004. — Vol. 13, No. 1. — P. 6–41.

- 46) Bujnowska, A. Access to European Statistical System Microdata / A. Bujnowska // In: *Data-Driven Policy Impact Evaluation*, ed. Crato N., Paruolo P. — Cham : Springer International Publishing, 2019, — P. 87–99.
- 47) Cantwell, J. Location and the multinational enterprise / J. Cantwell // *Journal of international business studies*. — 2009. — Vol. 40, No. 1. — P. 35–41.
- 48) Carpenter, C. W. Empirical methods in business location research / C. W. Carpenter, A. Van Sandt, S. Loveridge // *Regional Studies, Regional Science*. — 2021. — Vol. 8, No. 1. — P. 344–361.
- 49) Casali, Y. Machine learning for spatial analyses in urban areas: a scoping review / Y. Casali, N. Y. Aydin, T. Comes // *Sustainable Cities and Society*. — 2022. — Vol. 85. — P. 104050.
- 50) Chen, L. Manufacturing facility location and sustainability: A literature review and research agenda / L. Chen, J. Olhager, O. Tang // *International Journal of Production Economics*. — 2014. — Vol. 149. — P. 154–163.
- 51) Christaller, W. *Central places in southern Germany* / W. Christaller. — Prentice Hall, 1966.
- 52) Cieřlik, A. Location Determinants of Japanese Multinationals in Poland: Do Special Economic Zones Really Matter for Investment Decisions? / A. Cieřlik, M. Ryan // *Journal of Economic Integration*. — 2005.— P. 475–496.
- 53) Coeurderoy, R. Regulatory environments and the location decision: Evidence from the early foreign market entries of new-technology-based firms / R. Coeurderoy, G. Murray // *Journal of International Business Studies*. — 2008. — Vol. 39. — P. 670–687.
- 54) Conte, A. Mapping innovative activity using microdata / A. Conte // *Applied Economics Letters*. — 2009. — Vol. 16, No. 18. — P. 1795–1799.
- 55) Crabbé, K. Taxes, agglomeration rents and location decisions of firms / K. Crabbé, K. De Bruyne // *De Economist*. — 2013. — Vol. 161, No. 4. — P. 421–446.
- 56) Cyert, R. Behavioral theory of the firm / R. Cyert, J. March // *Organizational Behavior 2*. — Routledge, 2015. — P. 60–77.

- 57) Dabney, D. Y. Do enterprise zone incentives affect business location decisions? / D. Y. Dabney // *Economic Development Quarterly*. — 1991. — Vol. 5, No. 4. — P. 325–334.
- 58) Davis, J. C. Regional Analysis Using Census Bureau Microdata at the Center for Economic Studies / J. C. Davis, B. P. Holly // *International Regional Science Review*. — 2006. — Vol. 29, No. 3. — P. 278–296.
- 59) Dembour, C. Competition for business location: A survey / C. Dembour // *Journal of Industry, Competition and Trade*. — 2008. — Vol. 8. — P. 89–111.
- 60) Díez-Vial, I. How knowledge links with universities may foster innovation: The case of a science park / I. Díez-Vial, A. Montoro-Sánchez // *Technovation*. — 2016. — Vol. 50. — P. 41–52.
- 61) Donaldson, D. The View from Above: Applications of Satellite Data in Economics / D. Donaldson, A. Storeygard // *Journal of Economic Perspectives*. — 2016. — Vol. 30, No. 4. — P. 171–198.
- 62) Dorji, U. J. A Machine Learning Approach to Estimate Median Income levels of Sub-Districts in Thailand using Satellite and Geospatial Data / U. J. Dorji, A. Plangprasopchok, N. Surasvadi, C. Siripanpornchana // *In Proceedings of the ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery, 2019*. — P. 11–14.
- 63) Druilhe, C. Emergence and growth of high-tech activity in Cambridge and Grenoble / C. Druilhe, E. Garnsey // *Entrepreneurship & Regional Development*. — 2000. — Vol. 12, No. 2. — P. 163–177.
- 64) Dubé, J. Location theories and business location decision: A micro-spatial investigation in Canada / J. Dubé, C. Brunelle, D. Legros // *The Review of Regional Studies*. — 2016. — Vol. 46, No. 2. — P. 143–170.
- 65) Duncan, G. T. Enhancing Access to Microdata While Protecting Confidentiality: Prospects for the Future / G. T. Duncan, R. W. Pearson // *Statist. Sci.* — 1991. — Vol. 6, No. 3.

- 66) Dunning, J. H. Location and the multinational enterprise: a neglected factor? / J. H. Dunning // *Journal of international business studies*. — 1998. — Vol. 29, No. 1. — P. 45–66.
- 67) Dunstan, K. Evaluation of alternative data sources for population estimates / K. Dunstan, M. Ryan. — Wellington N.Z. : Statistics New Zealand, 2011. — P. 37.
- 68) Elhorst, J. P. Spatial econometrics: from cross-sectional data to spatial panels / J. P. Elhorst. — Berlin : Springer, 2014. — Vol. 479. — P. 480.
- 69) Elwood, S. Researching volunteered geographic information: Spatial data, geographic research, and new social practice / S. Elwood, M. F. Goodchild, D. Z. Sui // *Annals of the association of American geographers*. — 2012. — Vol. 102, No. 3. — P. 571–590.
- 70) Feldman, M. P. The geographic sources of innovation: technological infrastructure and product innovation in the United States / M. P. Feldman, R. Florida // *Annals of the association of American Geographers*. — 1994. — Vol. 84, No. 2. — P. 210–229.
- 71) Florida, R. Entrepreneurship, creativity, and regional economic growth / R. Florida // *The emergence of entrepreneurship policy*. — 2003. — P. 39–58.
- 72) Fujita, M. Thünen and the new economic geography / M. Fujita // *Regional Science and Urban Economics*. — 2012. — Vol. 42, No. 6. — P. 907–912.
- 73) Fujita, M. Economics of agglomeration / M. Fujita, J. F. Thisse // *Journal of the Japanese and international economies*. — 1996. — Vol. 10, No. 4. — P. 339–378.
- 74) Furman, J. L. Location and organizing strategy: Exploring the influence of location on the organization of pharmaceutical research / J. L. Furman // *In Geography and strategy*. — 2003. — Vol. 20. — P. 49–87.
- 75) Gabe, T. M. Tradeoffs between local taxes and government spending as determinants of business location / T. M. Gabe, K. P. Bell // *Journal of Regional Science*. — 2004. — Vol. 44, No. 1. — P. 21–41.
- 76) Galindo-Rueda, F. OECD taxonomy of economic activities based on R&D intensity / F. Galindo-Rueda, F. Verger // *OECD Science, Technology and Industry Working Papers*. — 2016. — No. 4.

- 77) García-Canal, E. Risk and the strategy of foreign location choice in regulated industries / E. García-Canal, M. F. Guillén // *Strategic Management Journal*. — 2008. — Vol. 29, No. 10. — P. 1097–1115.
- 78) Goerzen, A. Business micro-location: factors, preferences and indicators to assess the influence factors / A. Goerzen, C. G. Asmussen, B. B. Nielsen // *Management and Production Engineering Review*. — 2013. — Vol. 4, No. 2. — P. 25–36.
- 79) Goerzen, A. Global cities and multinational enterprise location strategy / A. Goerzen, C. G. Asmussen, B. B. Nielsen // *Journal of international business studies*. — 2013. — Vol. 44, No. 5. — P. 427–450.
- 80) Gothwal, S. Plant location selection of a manufacturing industry using analytic hierarchy process approach / S. Gothwal, R. Saha // *International Journal of Services and Operations Management*. — 2015. — Vol. 22, No. 2. — P. 235–255.
- 81) Haeussler, C. How patenting informs VC investors — The case of biotechnology / C. Haeussler, D. Harhoff, E. Mueller // *Research Policy*. — 2014. — Vol. 43, No. 8. — P. 1286–1298.
- 82) Hill, J. The effect of location on the performance of high technology firms / J. Hill, J. L. Naroff // *Financial Management*. — 1984. — P. 27–36.
- 83) Hoover, E. M. Location of economic activity / E. M. Hoover. — New York : McGraw-Hill Book Company, Inc., 1948.
- 84) Hsu, W. K. Earthquake risk assessment and optimal risk management strategies for hi-tech fabs in Taiwan / W. K. Hsu, W. L. Chiang, C. W. Chen // *Natural hazards*. — 2013. — Vol. 65. — P. 2063–2076.
- 85) Isard, W. Location and space-economy: a general theory relating to industrial location, market areas, land use, trade, and urban structure / W. Isard. — 1965. — No. 338 I8. — P. 350
- 86) Izadikhah, M. A new preference voting method for sustainable location planning using geographic information system and data envelopment analysis / M. Izadikhah, R. F. Saen // *Journal of Cleaner Production*. — 2016.

- 87) Jdanov, D. A. The short-term mortality fluctuation data series, monitoring mortality shocks across time and space / D. A. Jdanov et al. // *Sci Data*. — 2021. — Vol. 8, No. 1. — P. 235.
- 88) Jing, C. Geospatial dashboards for monitoring smart city performance / C. Jing, M. Du, S. Li, S. Liu // *Sustainability*. — 2019. — Vol. 11, No. 20. — P. 1–23.
- 89) Johansson, B. Innovation activities explained by firm attributes and location / B. Johansson, H. Lööf // *Econ. Innov. New Techn.* — 2008. — Vol. 17, No. 6. — P. 533–552.
- 90) Johnson, J. D. The role of economic and quality of life values in rural business location / J. D. Johnson, R. Rasker // *Journal of Rural Studies*. — 1995. — Vol. 11, No. 4. — P. 405–416.
- 91) Karakaya, F. Underlying dimensions of business location decisions / F. Karakaya, C. Canel. — *Industrial management & data systems*, 1998.
- 92) Kaufmann, D. The role of location and regional networks for biotechnology firms in Israel / D. Kaufmann, D. Schwartz, A. Frenkel, D. Shefer // *European Planning Studies*. — 2003. — Vol. 11, No. 7. — P. 823–840.
- 93) Kenney, M. Supporting the High-Technology Entrepreneur: Support Network Geographies for Semiconductor, Telecommunications Equipment, and Biotechnology Start-ups / M. Kenney, D. Patton. — UC Berkeley : Berkeley Roundtable on the International Economy, 2004.
- 94) Kilvits, K. Living environment as a location decision factor for manufacturing enterprises / K. Kilvits // *China-USA Business Review*. — 2012. — Vol. 11, No. 2. — P. 217–224
- 95) Kimelberg, S. M. Evaluating the importance of business location factors: The influence of facility type / S. M. Kimelberg, E. Williams // *Growth and Change*. — 2013. — Vol. 44, No. 1. — P. 92–117.
- 96) Kinne, J. Analyzing and predicting micro-location patterns of software firms / J. Kinne, B. Resch // *ISPRS International Journal of Geo-Information*. — 2018. — Vol. 7, No. 1. — P. 1.

- 97) Koshovets, O. B. Financial performance and technological modernization of Russian Hi-technology complex and their role to boost economic growth / O. B. Koshovets // *Journal of International Scientific Publications: Economy & Business*. — 2014. — Vol. 8, No. 2. — P. 633–340.
- 98) Kozyreva, P. Data Resource Profile: The Russia Longitudinal Monitoring Survey — Higher School of Economics (RLMS–HSE) Phase II: Monitoring the Economic and Health Situation in Russia, 1994–2013 / P. Kozyreva, M. Kosolapov, B. M. Popkin // *International Journal of Epidemiology*. — 2016. — Vol. 45, No. 2. — P. 395–401.
- 99) Krugman, P. The role of geography in development / P. Krugman // *International regional science review*. — 1999. — Vol. 22, No. 2. — P. 142–161.
- 100) Krugman, P. Urban Concentration: The Role of Increasing Returns and Transport Costs / P. Krugman // *International Regional Science Review*. — 1996. — Vol. 19, No. 1-2. — P. 5–30.
- 101) Kusumoto, S. I. Leontief Technology and the Location of the Firm in a Weber Triangle — Specific Localization Theorems / S. I. Kusumoto // *Journal of Regional Science*. — 1985. — Vol. 25, No. 3. — P. 443–451.
- 102) Lavallin, A. Machine learning in geography — Past, present, and future / A. Lavallin, J. A. Downs // *Geography Compass*. — 2021. — Vol. 15, No. 5
- 103) Lee, Y. Geographic redistribution of US manufacturing and the role of state development policy / Y. Lee // *Journal of Urban Economics*. — 2008. — Vol. 64, No. 2. — P. 436–450.
- 104) Lehnert, P. Proxying economic activity with daytime satellite imagery: Filling data gaps across time and space / P. Lehnert et al. // *PNAS Nexus*. — 2023. — Vol. 2, No. 4.
- 105) Li, N. Location Types of Enterprise R&D Institutions in Shanghai / N. Li // *Geographical review of Japan series B*. — 2012. — Vol. 84, No. 2. — P. 44–59.
- 106) Li, W. How does machine learning compare to conventional econometrics for transport data sets? A test of ML versus MLE / W. Li, K. M. Kockelman // *Growth and Change*. — 2022. — Vol. 53, No. 1. — P. 342–376.

- 107) Libório, M. P. Estimating municipal economic activity: An alternative data-based approach / M. P. Libório et al. // *Remote Sensing Applications: Society and Environment*. — 2022. — Vol. 28.
- 108) Libório, M. P. Economic analysis through alternative data and big data techniques: what do they tell about Brazil? / M. P. Libório, P. I. Ekel, C. A. P. Da Silva Martins // *SN Bus Econ*. — 2022. — Vol. 3, No. 1. — P. 3.
- 109) Löfsten, H. Science Parks and the growth of new technology-based firms — academic-industry links, innovation and markets / H. Löfsten, P. Lindelöf // *Research policy*. — 2002. — Vol. 31, No. 6. — P. 859–876.
- 110) Lola, I. Measurement of digital activity in medium, high-tech and low-tech manufacturing industries / I. Lola, M. Bakeev // *Higher School of Economics Research Paper No. WP BRP*, 2019. — Vol. 95.
- 111) Losch, A. *Economics of location* / A. Losch. — 1954.
- 112) Malecki, E. J. Industrial location and corporate organization in high technology industries / E. J. Malecki // *Economic Geography*. — 1985. — Vol. 61, No. 4. — P. 345–369.
- 113) Malik, S. *Enterprise dashboards: design and best practices for IT* / S. Malik. — John Wiley & Sons, 2005. — P. 240
- 114) Malmberg, A. *Industrial geography: location and learning* / A. Malmberg // *Progress in human geography*. — 1997. — Vol. 21, No. 4. — P. 573–582.
- 115) Markusen, A. R. *The rise of the gunbelt: The military remapping of industrial America* / F. R. Markusen. — Oxford University Press, USA, 1991.
- 116) Marrocu, E. Proximity, networking and knowledge production in Europe: What lessons for innovation policy? / E. Marrocu, R. Paci, S. Usai // *Technological Forecasting and Social Change*. — 2013. — Vol. 80, No. 8. — P. 1484–1498.
- 117) Marshall, A. *Principles of Economics: Text* / A. Marshall. — London; New York : Macmillan for the Royal Economic Society, 1961. — Vol. 1
- 118) Maximenko, D. GIS for location planning of banks' physical networks / D. Maximenko, M. Maximenko // *Regional Studies, Regional Science*. — 2021. — Vol. 8, No. 1. — P. 362–365.

- 119) McCann, B. T. Location matters: where we have been and where we might go in agglomeration research / B. T. McCann, T. B. Folta // *Journal of management*. — 2008. — Vol. 34, No. 3. — P. 532–565.
- 120) McCann, P. *Industrial location economics* / P. McCann. — Edward Elgar Publishing, 2002.
- 121) McCann, P. Location, agglomeration and infrastructure / P. McCann, D. Shefer // *Fifty Years of Regional Science*. — 2004. — P. 177–196.
- 122) McGuckin, R. H. *Establishment Microdata for Economic Research and Policy Analysis: Looking Beyond the Aggregates* / R. H. McGuckin // *Journal of Business & Economic Statistics*. — 1995. — Vol. 13, No. 1. — P. 12–126.
- 123) Melo, M. T. Facility location and supply chain management — A review / M. T. Melo, S. Nickel, F. Saldanha-Da-Gama // *European journal of operational research*. — 2009. — Vol. 196, No. 2. — P. 401–412.
- 124) Midelfart-Knarvik, K. H. *The location of European industry* / K. H. Midelfart-Knarvik, H. G. Overman, S. J. Redding, A. J. Venables. — 2000.
- 125) Mikhaylov A., Mikhaylova A., Maximenko D., Maximenko M., Hvalej D. Coastal regions in the geography of innovation activity: a comparative assessment of marine basins // *Geographica Pannonica*. 2022. Vol. 26, No. 4. P. 345–355.
- 126) Mikhaylova A., Maximenko D., Hvalej D., Maximenko M., Mikhaylov A. Geoinformation representation of maritime knowledge flows: new frontiers of coastalization // *IOP Conference Series: Earth and Environmental Science*. 2022. Vol. 1087. Article 012038.
- 127) Morgan, K. The exaggerated death of geography: learning, proximity and territorial innovation systems / K. Morgan // *Journal of Economic Geography*. — 2004. — Vol. 4, No. 1. — P. 3–21.
- 128) Neumann, P. J. *Toward Better Data Dashboards for US Drug Value Assessments* / P. J. Neumann // *Value in Health*. — 2021. — Vol. 24, No. 10. — P. 1484–1489.

- 129) Parikh, S. Data Analytics and Data Visualization for the Pharmaceutical Industry / S. Parikh, R. Patel, D. Khunt, V. P. Chavda, L. Vora // *Bioinformatics Tools for Pharmaceutical Drug Product Development*. — 2023. — P. 55–76.
- 130) Pauwels, K. Dashboards as a service: why, what, how, and what research is needed? / K. Pauwels, T. Ambler, B. H. Clark, P. LaPointe, D. Reibstein, B. Skiera et al. // *Journal of Service Research*. — 2009. — Vol. 12, No. 2. — P. 175–189.
- 131) Porter, M. E. *On competition* / M. E. Porter. — Harvard Business Press, 2008.
- 132) Psyllidis, A. Regionalization of social interactions and points-of-interest location prediction with geosocial data / A. Psyllidis, J. Yang, A. Bozzon // *IEEE Access*. — 2018. — Vol. 6. — P. 34334–34353.
- 133) Rammer, C. Microgeography of innovation in the city: Location patterns of innovative firms in Berlin / C. Rammer, J. Kinne, K. Blind // *ZEW-Centre for European Economic Research Discussion Paper*. — 2016. — No. 16–080.
- 134) Reades, J. Understanding urban gentrification through machine learning / J. Reades, J. De Souza, P. Hubbard // *Urban Studies*. — 2019. — Vol. 56, No. 5. — P. 922–942.
- 135) Rilla, N. R&D (re) location and offshore outsourcing: a management perspective / N. Rilla, M. Squicciarini // *International Journal of Management Reviews*. — 2011. — Vol. 13, No. 4. — P. 393–413.
- 136) Rodrigue, J. P. *The geography of transport systems* / J. P. Rodrigue. — Taylor & Francis, 2016.
- 137) Rogers, E. M. *Diffusion of innovations* / E. M. Rogers. — Simon and Schuster, 2010.
- 138) Salvesen, D. The importance of quality of life in the location decisions of new economy firms / D. Salvesen, H. Renski // *Report of Center for Urban and Regional Studies of University of North Carolina at Chapel Hill*, 2003.
- 139) Santos, J. D. F. Business micro-location: factors, preferences and indicators to assess the influence factors / J. D. F. Santos // *Management and Production Engineering Review*. — 2013. — P. 25–36.

- 140) Shearmur, R. Intrametropolitan patterns of high-order business service location: A comparative study of seventeen sectors in Ile-de-France / R. Shearmur, C. Alvergne // *Urban Studies*. — 2002. — Vol. 39, No. 7. — P. 1143–1163.
- 141) Simon, H. A. A behavioral model of rational choice / H. A. Simon // *The quarterly journal of economics*. — 1955. — (No.?) — P. 99–118.
- 142) Staber, U. Spatial Proximity and Firm Survival in a Declining Industrial District: The Case of Knitwear Firms in Baden-Württemberg / U. Staber // *Regional Studies*. — 2001. — Vol. 35, No. 4. — P. 329–341.
- 143) Stuart, T. The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms / T. Stuart, O. Sorenson // *Research policy*. — 2003. — Vol. 32, No. 2. — P. 229–253.
- 144) Thornhill, S. Knowledge, innovation and firm performance in high-and low-technology regimes / S. Thornhill // *Journal of business venturing*. — 2006. — Vol. 21, No. 5. — P. 687–703.
- 145) Van Der Panne, G. The odd role of proximity in knowledge relations: high-tech in the Netherlands / G. Van Der Panne, W. Dolfsma // *Tijdschrift voor economische en sociale geografie*. — 2003. — Vol. 94, No. 4. — P. 453–462.
- 146) Vlachou, C. The evolution of studies on business location factors / C. Vlachou, O. Iakovidou // *Journal of Developmental Entrepreneurship*. — 2015. — Vol. 20, No. 04.
- 147) Von Thünen, J. H. *Isolated state: an English edition of Der isolierte Staat* / J. H. Von Thünen. — Pergamon Press, 1966.
- 148) Wallerstein, I. The time of space and the space of time: The future of social science / I. Wallerstein // *Political Geography*. — 1998. — Vol. 17, No. 1. — P. 71–82.
- 149) Weber, A. *Theory of the Location of Industries* / A. Weber. — University of Chicago Press, 1929.
- 150) Witkowski, J. The macro-and micro-environmental factors of decisions of production facility location by Japanese companies in Poland / J. Witkowski, K.

- Cheba, M. Kiba-Janiak // In Forum Scientiae Oeconomia. — 2017. — Vol. 5, No. 1. — P. 43–56.
- 151) Wojan, T. R. Location patterns of high growth industries in rural counties / T. R. Wojan, G. C. Pulver // Growth and Change. — 1995. — Vol. 26, No. 1. — P. 3–22.
- 152) Wolf, L. Learning Geographical Manifolds: A Kernel Trick for Geographical Machine Learning / L. Wolf, E. Knaap. — 2019.
- 153) Woodward, D. Beyond the Silicon Valley: University R&D and high-technology location / D. Woodward, O. Figueiredo, P. Guimaraes // Journal of Urban Economics. — 2006. — Vol. 60, No. 1. — P. 15–32.
- 154) Wu, J. Economic Transition and Changing Location of Manufacturing Industry in China: A Study of the Yangtze River Delta / J. Wu, Y. D. Wei, Q. Li, F. Yuan. — Sustainability, 2018.
- 155) Xia, T. Foreign competition, domestic competition and innovation in Chinese private high-tech new ventures / T. Xia, X. Liu // Journal of International Business Studies. — 2017. — Vol. 48. — P. 716–739.
- 156) Yang, J. An AHP decision model for facility location selection / J. Yang, H. Lee. — Facilities, 1997.
- 157) Yesudas, M. Intelligent operational dashboards for smarter commerce using big data / M. Yesudas, G. Menon, V. Ramamurthy // IBM Journal of Research and Development. — 2014. — Vol. 58, No. 5/6.
- 158) Zhuang, L. Changing imbalance: Spatial production of national high-tech industrial development zones in China (1988–2018) / L. Zhuang, C. Ye // Land Use Policy. — 2020. — Vol. 94. — P. 104512.
- 159) SAS Blogs. How to visualize a kernel density estimate: <https://blogs.sas.com/content/iml/2016/07/27/visualize-kernel-density-estimate.html> (дата обращения 11.12.2023)
- 160) Yandex Datalens: <https://cloud.yandex.ru/ru/docs/datalens/> (дата обращения 11.12.2023)

ПРИЛОЖЕНИЕ

Приложение 1. Блокнот с ходом машинного обучения для решения задачи размещения высокотехнологичного бизнеса

Модель оценки потенциала территории для размещения высокотехнологичного бизнеса

На первом этапе загрузим необходимые библиотеки.

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px

from catboost import CatBoostClassifier, CatBoostRegressor, Pool, cv

import lightgbm as lgb

from sklearn.model_selection import train_test_split
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import cross_val_score

from sklearn.tree import DecisionTreeRegressor

from sklearn.ensemble import RandomForestRegressor

from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score

from sklearn.dummy import DummyRegressor

from sklearn.pipeline import Pipeline

from sklearn.decomposition import PCA

from sklearn.preprocessing import StandardScaler

import shap

import warnings
warnings.filterwarnings("ignore")

from IPython.core.display import HTML
HTML("""
<style>
.output_png {
    display: table-cell;
    text-align: center;
    vertical-align: middle;
}
</style>
""")
```

Using `tqdm.autonotebook.tqdm` in notebook mode. Use `tqdm.tqdm` instead to force console mode (e.g. in jupyter console)

Out[1]:

Загрузка и предобработка данных

Загрузим данные для моделирования из заранее подготовленной таблицы, основанной на

геопривязке факторов размещения к предприятиям высокотехнологического комплекса.

```
In [2]: df = pd.read_excel(r"C:\Users\dmaximenko\Desktop\Model\model_data_c.xlsx") #директория
```

```
In [3]: data = df.fillna(0)
```

Исключим Москву и Санкт-Петербург из периметра на основе предварительного анализа работы модели.

```
In [4]: data = data.set_index('fid')
data = data.query('moscow != 1')
data = data.query('spb != 1')
data = data.drop(['spb', 'moscow'], axis = 1)
data.head(2)
```

```
Out[4]:
```

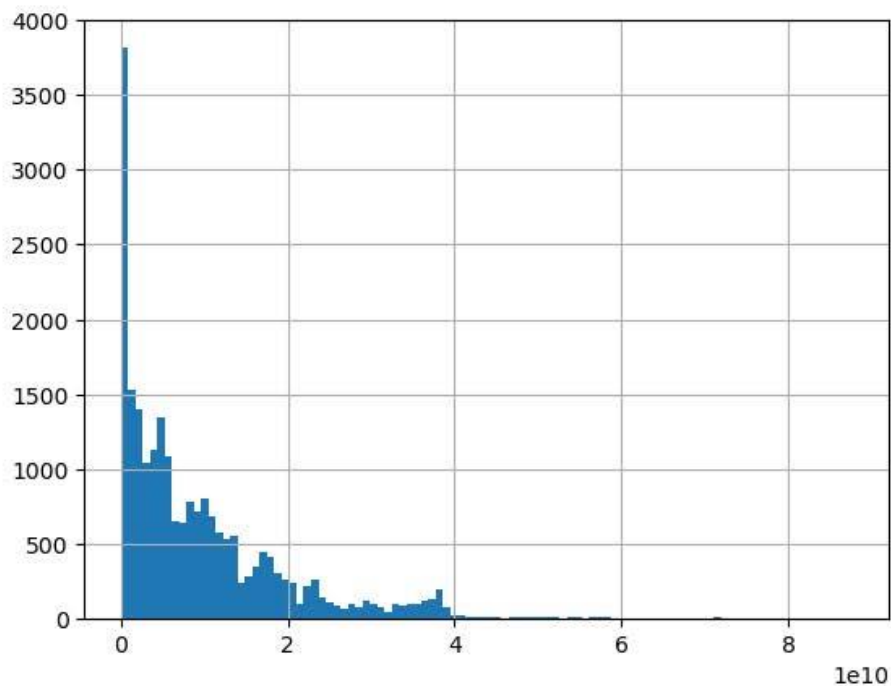
	okved	population	inc_kernel	coast	border	dim_jun	clim_dec	distance_reg	distance_mil	distance_
fid										
2.0	29.0	1425508.0	7.616018e+09	0.0	0.0	16.924566	-13.744932	20.180966	16.203067	10.068
3.0	28.0	1104648.0	2.335375e+10	0.0	0.0	18.099254	-10.520304	0.589425	2.273729	0.658

2 rows × 28 columns

Прологарифмируем зависимую переменную для упрощения работы алгоритмов машинного обучения

```
In [5]: data['inc_kernel'].hist(bins=100)
```

```
Out[5]: <AxesSubplot:>
```



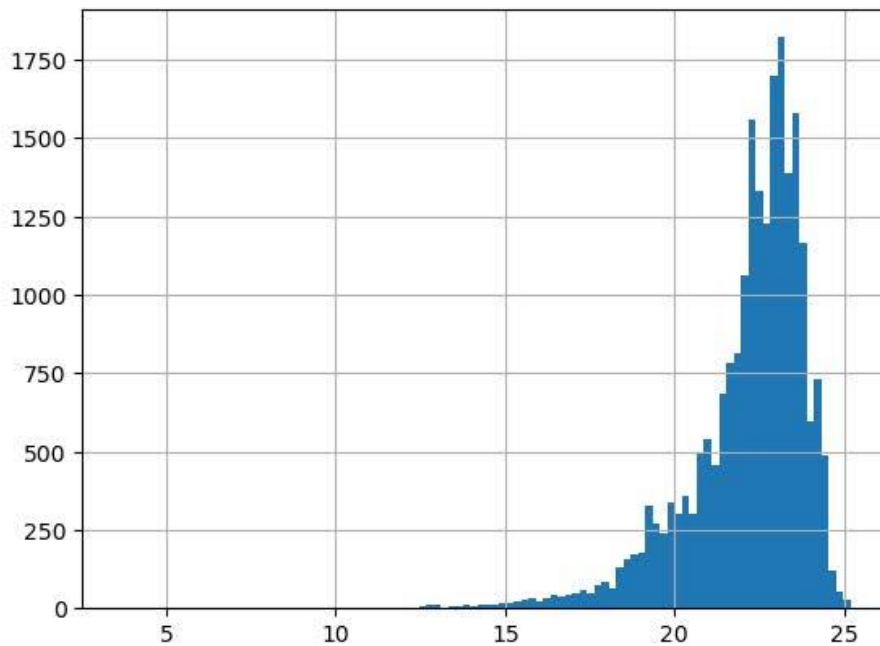
```
In [6]: data['inc_kernel_ln'] = np.log(data['inc_kernel'])
```



```
data = data.drop(['inc_kernel'], axis = 1)
```

```
In [7]: data = data.query('inc_kernel_ln > 0')
data['inc_kernel_ln'].hist(bins=100)
```

```
Out[7]: <AxesSubplot:>
```

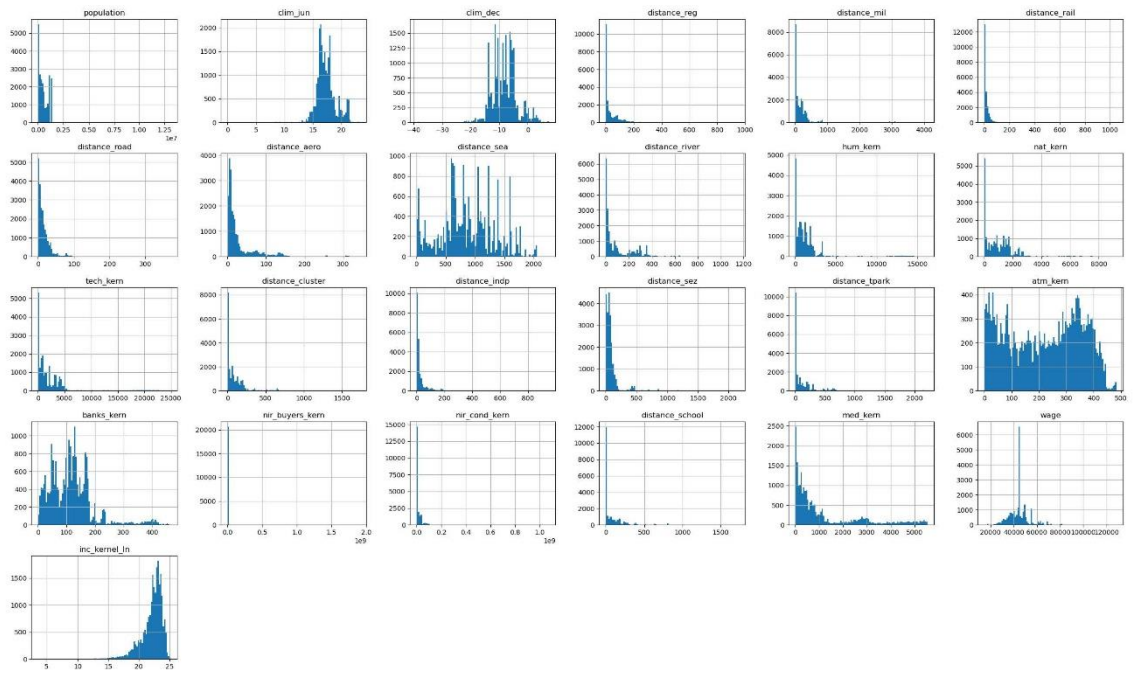


Выделим категориальные переменные

```
In [8]: data['coast'] = data['coast'].astype("category")
data['border'] = data['border'].astype("category")
data['okved'] = data['okved'].astype("category")
```

На основе анализа гистограмм распределения независимых переменных также проведем логарифмирование факторов закупок и исполнения НИР.

```
In [9]: hist = data.hist(figsize = (30,40), layout = (11,6), bins = 100)
```

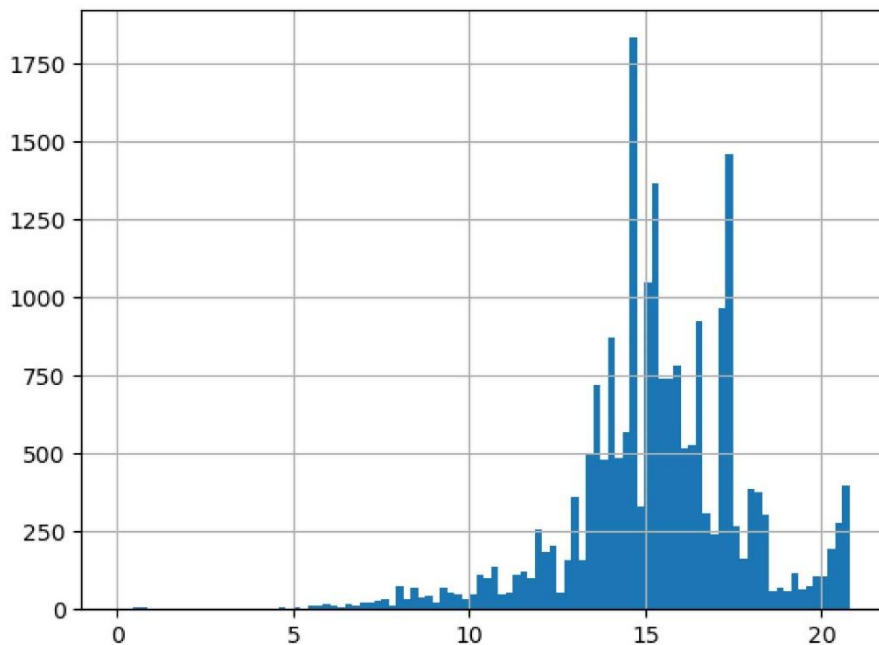


```
In [10]: data['nir_buyers_ln'] = np.log(data['nir_buyers_kern'])
data = data.drop(['nir_buyers_kern'], axis = 1)
data['nir_cond_ln'] = np.log(data['nir_cond_kern'])
data = data.drop(['nir_cond_kern'], axis = 1)
```

```
In [11]: data = data.query('nir_buyers_ln > 0')
data = data.query('nir_cond_ln > 0')
```

```
In [12]: data['nir_cond_ln'].hist(bins = 100)
```

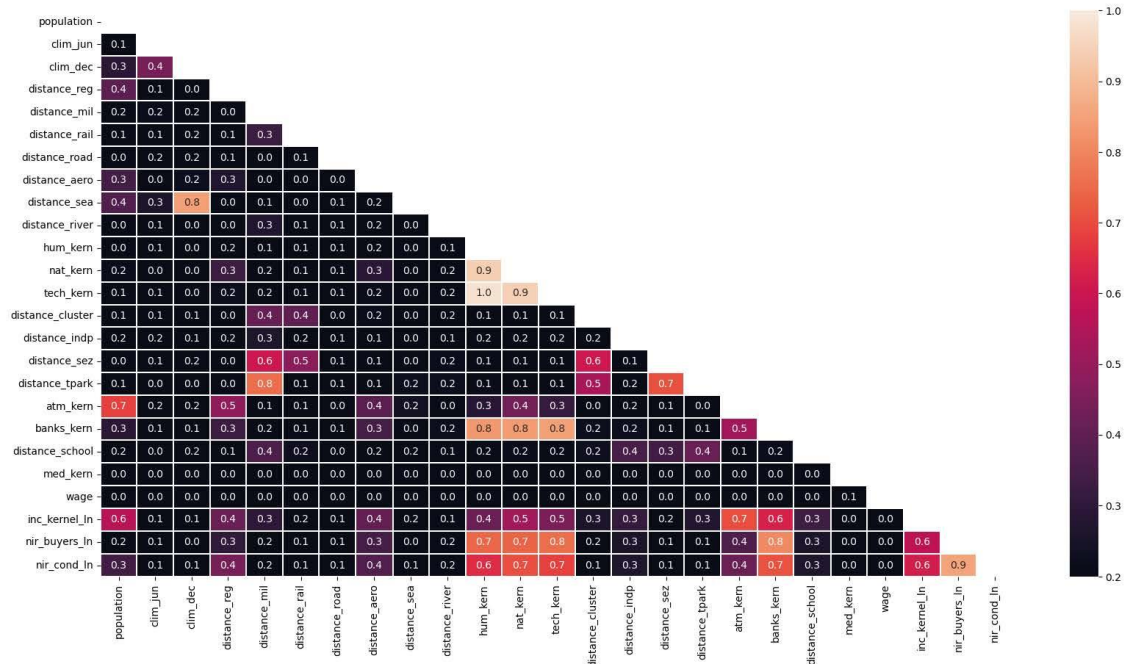
Out[12]: <AxesSubplot:>



Проведем корреляционный анализ. Факторы, связанные с выпуском специалистами вузов разных специальностей и исполнителей и заказчиков НИР имеют крайне высокую корреляцию, поэтому заменим их на объединенные переменные.

```
In [13]: corr = abs(data.corr())
mask = np.zeros_like(corr)
mask[np.triu_indices_from(mask)] = True
c, ax = plt.subplots(figsize=(20, 10))
sns.heatmap(corr, mask = mask, vmin=0.2, vmax=1, annot = True, fmt=".1f", linewidths=.1,
```

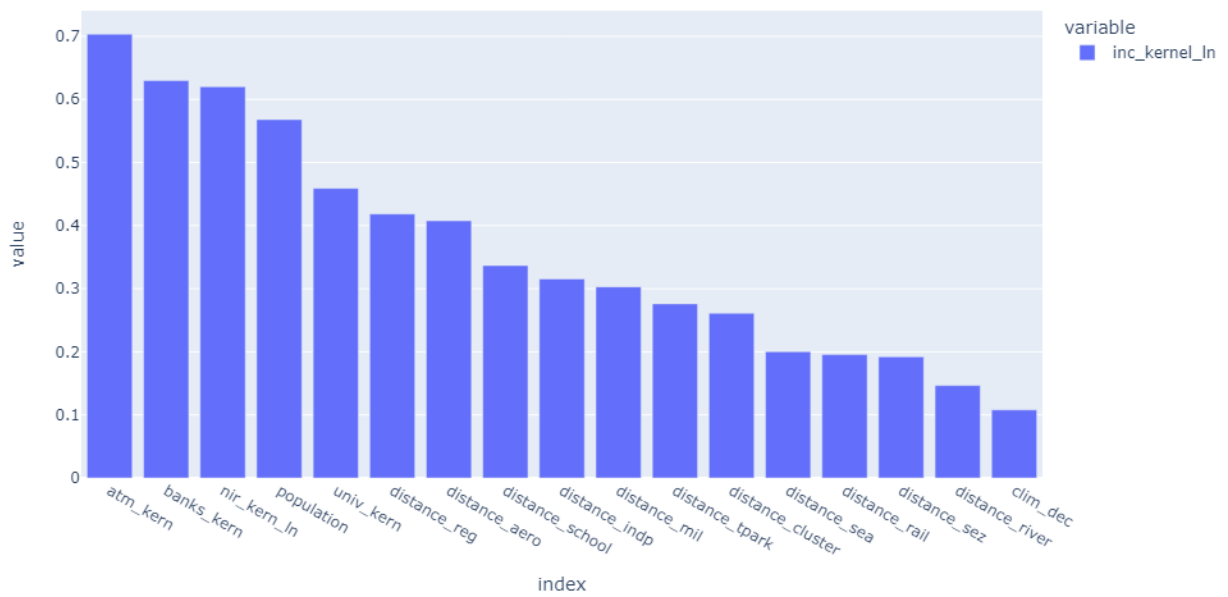
Out[13]: <AxesSubplot:>



```
In [14]: data['univ_kern'] = data['hum_kern'] + data['nat_kern'] + data['tech_kern']
data['nir_kern_ln'] = data['nir_buyers_ln'] + data['nir_cond_ln']
data = data.drop(['hum_kern', 'nat_kern', 'tech_kern', 'nir_buyers_ln', 'nir_cond_ln'], a
```

```
In [15]: corr = abs(data.corr())
```

```
In [16]: x = abs(corr['inc_kernel_ln']).sort_values(ascending=False)
x = x[x<1]
fig = px.bar(x[0.1 < x])
fig.show()
```



Обработаем категориальные переменные методом OneHotEncoding (OHE)

```
In [17]: data_oh = pd.get_dummies(data, columns=['okved', 'coast', 'border'], drop_first=True)
data_oh.head(2)
```

```
Out[17]:
```

	population	clim_jun	clim_dec	distance_reg	distance_mil	distance_rail	distance_road	distance_aero	dist
fid									
2.0	1425508.0	16.924566	-13.744932	20.180966	16.203067	10.068908	10.532736	20.242879	16
3.0	1104648.0	18.099254	-10.520304	0.589425	2.273729	0.658450	13.715829	15.201860	13

2 rows x 29 columns

Моделирование

```
In [18]: features = data_oh.drop(['inc_kernel_ln'], axis=1)
target = data_oh['inc_kernel_ln']

features_train_f, features_test_f, target_train_f, target_test_f = train_test_split(
    data_oh.drop(['inc_kernel_ln'], axis=1), data_oh['inc_kernel_ln'], test_size=0.33,
```

```
In [19]: def rank(model_name):
a = model_name.cv_results_['mean_train_r2']
a = a[~np.isnan(a)]
b = model_name.cv_results_['mean_train_neg_root_mean_squared_error']
b = b[~np.isnan(b)]
c = model_name.cv_results_['mean_test_neg_root_mean_squared_error']
c = c[~np.isnan(c)]
print('Лучшие параметры:', model_name.best_params_)
print('Значение на тренировочной выборке: R2', a.mean().round(2), '/ RMSE',
      -b.mean().round(2))
print('Значение на валидационной выборке: R2 ', model_name.best_score_.round(2), '/ R',
      -c.mean().round(2))
```

```
In [20]: dummy = DummyRegressor(strategy='median')
pca = PCA(random_state=43)
scaler = StandardScaler()
```

```
dc = DecisionTreeRegressor(random_state=43)
rf = RandomForestRegressor(random_state=43)
lgbx = lgb.LGBMRegressor(verbose=-1, random_state=43)
cbx = CatBoostRegressor(random_state=43, verbose = False)
```

```
In [21]: pipe = Pipeline(steps=[("scaler", scaler), ("pca", pca), ("model", dummy)])

parameters = {"scaler": [None, scaler],
              "pca__n_components": [None, 3, 5],
              }

dummy_f = GridSearchCV(pipe,scoring = ("r2", 'neg_root_mean_squared_error'), refit = 'r2
                    param_grid=parameters, return_train_score=True).fit(features_train_f,
                    target_train_f)

best_dummy_f = dummy_f.best_estimator_

best_dummy_f.fit(features_train_f, target_train_f)
```

```
Out[21]: Pipeline(steps=[('scaler', None), ('pca', PCA(random_state=43)),
                        ('model', DummyRegressor(strategy='median'))])
```

```
In [22]: rank(dummy_f)
```

```
Лучшие параметры: {'pca__n_components': None, 'scaler': None}
Значение на тренировочной выборке: R2 -0.05 / RMSE 1.62
Значение на валидационной выборке: R2 -0.05 / RMSE 1.62
```

```
In [23]: pipe = Pipeline(steps=[("scaler", scaler), ("pca", pca), ("model", dc)])

parameters = {"scaler": [None, scaler],
              "pca__n_components": [None, 3, 5],
              'model__max_depth': range(1, 20),
              'model__max_features': ('auto', 'sqrt', 'log2')
              }

dc_tree_f = GridSearchCV(pipe,scoring = ("r2", 'neg_root_mean_squared_error'), refit = '
                    param_grid=parameters, return_train_score=True).fit(features_train_f,
                    target_train_f)

best_tree_f = dc_tree_f.best_estimator_

best_tree_f.fit(features_train_f, target_train_f)
```

```
Out[23]: Pipeline(steps=[('scaler', None), ('pca', PCA(random_state=43)),
                        ('model',
                         DecisionTreeRegressor(max_depth=15, max_features='auto',
                                                random_state=43))])
```

```
In [33]: rank(dc_tree_f)
```

```
Лучшие параметры: {'model__max_depth': 15, 'model__max_features': 'auto', 'pca__n_components': None, 'scaler': None}
Значение на тренировочной выборке: R2 0.8 / RMSE 0.62
Значение на валидационной выборке: R2 0.89 / RMSE 0.85
```

```
In [25]: predictions = dc_tree_f.predict(features_test_f)
```

```
In [26]: pipe = Pipeline(steps=[("scaler", scaler), ("pca", pca), ("model", rf)])

parameters = {"scaler": [None, scaler],
              "pca__n_components": [None, 3, 5],
              'model__n_estimators': range(1, 10),
              'model__max_depth': range(1,10)
              }
```

```
rf_tree_f = GridSearchCV(pipe,scoring = ("r2", 'neg_root_mean_squared_error'), refit = '
    param_grid=parameters, return_train_score=True).fit(features_train_f, target_train_f)

best_rf_f = rf_tree_f.best_estimator_

best_rf_f.fit(features_train_f, target_train_f)
```

```
Out[26]: Pipeline(steps=[('scaler', None), ('pca', PCA(random_state=43)),
    ('model',
    RandomForestRegressor(max_depth=9, n_estimators=9,
    random_state=43))])
```

```
In [27]: rank(rf_tree_f)
```

```
Лучшие параметры: {'model__max_depth': 9, 'model__n_estimators': 9, 'pca__n_components':
None, 'scaler': None}
Значение на тренировочной выборке: R2 0.71 / RMSE 0.82
Значение на валидационной выборке: R2 0.91 / RMSE 0.86
```

```
In [28]: pipe = Pipeline(steps=[("scaler", scaler), ("pca", pca), ("model", lgbr)])
```

```
parameters = {"scaler": [None, scaler],
    "pca__n_components": [None, 3, 5],
    "model__max_depth": range(0,10),
    "model__learning_rate" : [0.01,0.1]
    }
```

```
lg_f = GridSearchCV(pipe,scoring = ("r2", 'neg_root_mean_squared_error'), refit = 'r2',
    param_grid=parameters, return_train_score=True).fit(features_train_f, target_train_f)

best_lg_f = lg_f.best_estimator_
best_lg_f.fit(features_train_f, target_train_f)
```

```
Out[28]: Pipeline(steps=[('scaler', None), ('pca', PCA(random_state=43)),
    ('model',
    LGBMRegressor(max_depth=0, random_state=43, verbose=-1))])
```

```
In [29]: rank(lg_f)
```

```
Лучшие параметры: {'model__learning_rate': 0.1, 'model__max_depth': 0, 'pca__n_component
s': None, 'scaler': None}
Значение на тренировочной выборке: R2 0.73 / RMSE 0.78
Значение на валидационной выборке: R2 0.94 / RMSE 0.83
```

Для catboost не подходят данные, обработанные ONE, поэтому для него делаем свой тренировочный датасет.

```
In [30]: #features_train_f, features_test_f, target_train_f, target_test_f = train_test_split(
    # data_ohc.drop(['inc_kernel_ln'], axis=1), data_ohc['inc_kernel_ln'], test_size=0.33

features_train_cb, features_test_cb, target_train_cb, target_test_cb = train_test_split(
    data_ohc.drop(['inc_kernel_ln'], axis=1), data_ohc['inc_kernel_ln'], test_size=0.33,
```

```
In [35]: pipe = Pipeline(steps=[("scaler", scaler), ("pca", pca), ("model", cbx)])
```

```
parameters = {"scaler": [None, scaler],
    "pca__n_components": [None, 3, 5],
    "model__depth": range(0,5),
    "model__iterations": [500, 1000],
    "model__learning_rate": [0.01, 0.1]}
```

```
cb_f = GridSearchCV(pipe,scoring = ("r2", 'neg_root_mean_squared_error'), refit = 'r2',
```

```

param_grid=parameters, return_train_score=True).fit(features_train)

best_cb_f = cb_f.best_estimator_
best_cb_f.fit(features_train_cb, target_train_cb)

```

Out[35]: Pipeline(steps=[('scaler', None), ('pca', PCA(random_state=43)), ('model', <catboost.core.CatBoostRegressor object at 0x0000020010135160>)])

In [36]: rank(cb_f)

Лучшие параметры: {'model__depth': 4, 'model__iterations': 1000, 'model__learning_rate': 0.1, 'pca__n_components': None, 'scaler': None}
Значение на тренировочной выборке: R2 0.62 / RMSE 0.9
Значение на валидационной выборке: R2 0.95 / RMSE 0.93

In [53]: predictions_cb_f = cb_f.predict(features_test_f)

In [80]: print('R2 на тестовой выборке ', round(r2_score(target_test_f, predictions_cb_f), 2))
print('RMSE на тестовой выборке ', round(mean_squared_error(target_test_f, predictions_cb_f), 2))

R2 на тестовой выборке 0.94
RMSE на тестовой выборке 0.16

Оптимизация лучшей модели с помощью shap-values

С помощью библиотеки shap оценим влияние отдельных переменных на предсказательную возможность модели.

```

In [55]: explainer = shap.Explainer(cb_f.predict, features_train_cb)
shap_values = explainer(features_train_cb)

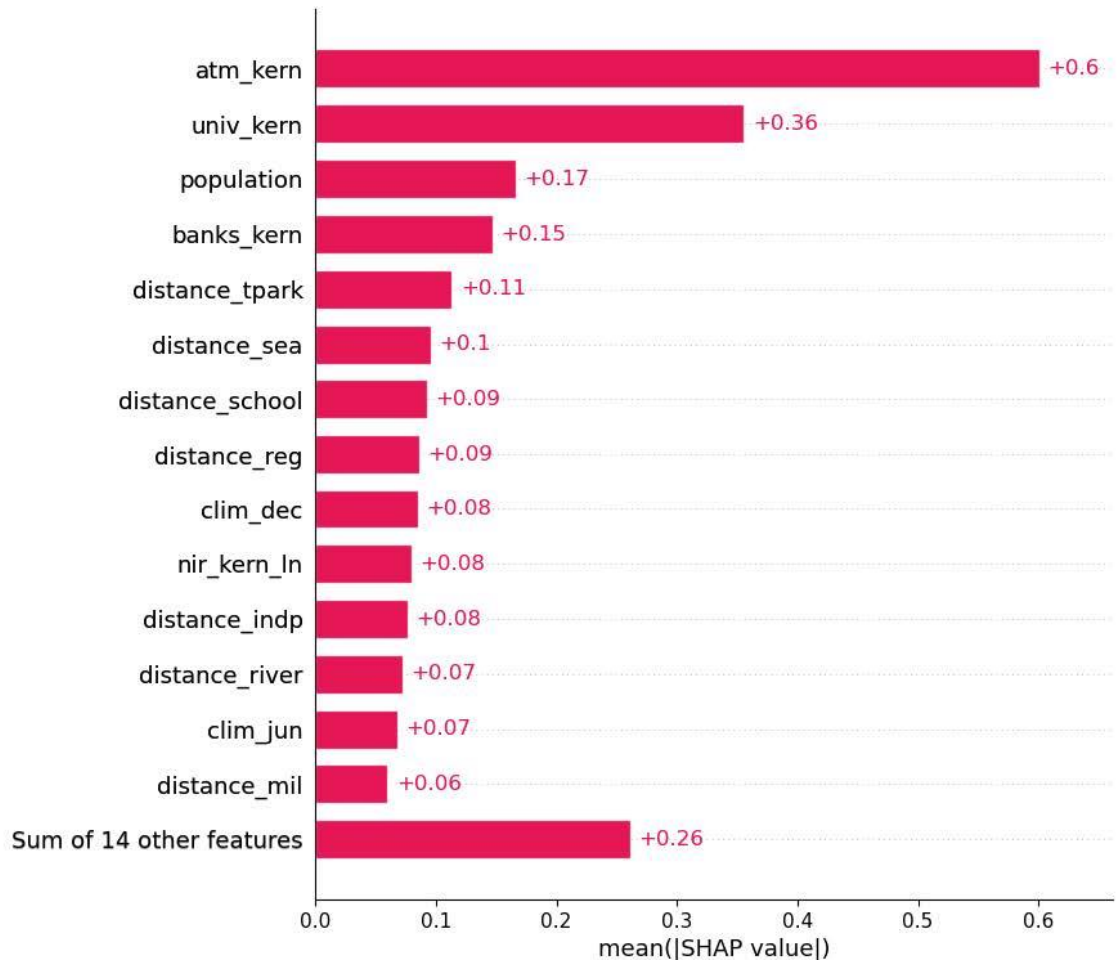
```

Permutation explainer: 14311it [1:03:42, 3.73it/s]

```

In [57]: shap.plots.bar(shap_values, max_display=15)

```



Из списка наиболее влиятельных показателей сформируем подвыборку для попытки оптимизации вычислений и работы модели.

```
In [87]: features_sc = data[['atm_kern', 'univ_kern', 'population', 'distance_tpark', 'distance_s
features_train_sc = features_train_f[['atm_kern', 'univ_kern', 'population', 'distance_t
target_train_sc = target_train_f

features_test_sc = features_test_f[['atm_kern', 'univ_kern', 'population', 'distance_tpa
target_test_sc = target_test_f
```

```
In [59]: pipe = Pipeline(steps=[("scaler", scaler), ("pca", pca), ("model", cbx)])

parameters = {"scaler": [None, scaler],
              "pca_n_components": [None, 3, 5],
              "model__depth": range(0, 5),
              "model__iterations": [500, 1000],
              "model__learning_rate": [0.01, 0.1]}

cb_sc = GridSearchCV(pipe, scoring = ("r2", 'neg_root_mean_squared_error'), refit = 'r2',
                    param_grid=parameters, return_train_score=True).fit(features_train_sc,
best_cb_sc = cb_sc.best_estimator_
best_cb_sc.fit(features_train_sc, target_train_sc)

Pipeline(steps=[('scaler', None), ('pca', PCA(random_state=43)),
```



```
Out[59]: ('model',
         <catboost.core.CatBoostRegressor object at 0x000002001A403DF0>)]
```

```
In [63]: rank(cb_sc)
```

```
Лучшие параметры: {'model__depth': 4, 'model__iterations': 1000, 'model__learning_rate':
0.1, 'pca__n_components': None, 'scaler': None}
Значение на тренировочной выборке: R2 0.65 / RMSE 0.85
Значение на валидационной выборке: R2 0.92 / RMSE 0.87
```

```
In [90]: predictions_cb_sc = cb_sc.predict(features_test_sc)
```

```
In [92]: print('R2 на тестовой выборке ', round(r2_score(target_test_sc,predictions_cb_sc),2))
print('RMSE на тестовой выборке ', round(mean_squared_error(target_test_sc,predictions_c
```

```
R2 на тестовой выборке 0.92
RMSE на тестовой выборке 0.2
```

Выгрузим оба предсказания для дальнейшей визуализации в ГИС.

```
In [96]: predictions_full_cb_f = pd.DataFrame(data = (cb_f.predict(features)))
predictions_full_cb_sc = pd.DataFrame(data = (cb_sc.predict(features_sc)))
```

```
In [97]: predictions_full_cb_f.to_excel(r"C:\Users\dmaksimenko\Desktop\Model\output1.xlsx")
predictions_full_cb_sc.to_excel(r"C:\Users\dmaksimenko\Desktop\Model\output2.xlsx")
target.to_excel(r"C:\Users\dmaksimenko\Desktop\Model\output3.xlsx")
```