

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Пермский национальный исследовательский политехнический
университет»

Факультет прикладной математики и механики
Направление 01.03.02 «Прикладная математика и
информатика»
Кафедра «Прикладная математика»

**Выпускная квалификационная работа
на соискание степени бакалавра**

На тему: «Многомерный статистический анализ состояния
полигонов твердых бытовых отходов»

Студент: Полежаева Елена Андреевна

**Руководитель дипломной
работы**
*Кандидат технических наук,
доцент кафедры ПМ*
Давыдов А.Р.

Пермь 2019 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	4
ГЛАВА 1. ОСНОВНЫЕ СВЕДЕНИЯ О ПОЛИГОНАХ ТБО.....	6
1.1. Твердые бытовые отходы и полигоны ТБО.....	6
1.2. Экологический мониторинг. Основные химико-биологические показатели.....	11
1.3. Рекультивация полигонов ТБО.....	14
1.4. Польза и выгода от ТБО.....	16
ГЛАВА 2. ПРИМЕНЯЕМЫЕ МЕТОДЫ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА.....	19
2.1. Корреляционно-регрессионный анализ.....	19
2.2. Кластерный анализ как метод группировки статистических данных.....	24
2.2.1. Иерархические методы кластерного анализа.....	27
2.2.2. Метод k -средних как итеративный метод кластерного анализа.....	30
2.3. Дисперсионный анализ как проверка качества кластеризации.....	32
2.4. Факторный анализ.....	34
ГЛАВА 3. АНАЛИЗ СТАБИЛЬНОСТИ ПОЛИГОНОВ ТБО.....	40
3.1. Исходные данные.....	40
3.2. Проведение кластерного анализа с целью определения стабильных объектов захоронения.....	42
3.3. Исследование изменения фракционного состава.....	52
3.4. Исследование изменения морфологического состава.....	58

3.5. Факторный анализ экологического мониторинга полигонов ТБО.....	60
ЗАКЛЮЧЕНИЕ.....	65
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	67
ПРИЛОЖЕНИЕ А.....	70
ПРИЛОЖЕНИЕ Б.....	70
ПРИЛОЖЕНИЕ В.....	71
ПРИЛОЖЕНИЕ Г.....	73
ПРИЛОЖЕНИЕ Д.....	75
ПРИЛОЖЕНИЕ Е.....	77

ВВЕДЕНИЕ

Одной из главных проблем современности является проблема отходов, а именно твердых бытовых отходов, наибольшей части продуктов потребления, которые потеряли свои качества, необходимые потребителю. На данный момент в России эта проблема превратилась в центральный вопрос, для решения которого с каждым годом вовлекается всё больше специалистов.

Крупные города являются главными источниками загрязнения. С каждым годом количество выбрасываемого мусора растет, и, соответственно, растут объёмы мусорных полигонов.

Существуют различные способы, с помощью которых можно из мусора получать выгоду, такие как использование переработанного мусора в качестве топлива, превращение свалочного газа в электроэнергию, добыча из углеродсодержащих отходов синтез-газа, из которого в дальнейшем создается синтетическое моторное топливо. Также участки земли, на которых размещаются такие полигоны, срок эксплуатации которых уже истёк, и все химико-биологические показатели пришли в стабильное состояние с окружающей средой, можно использовать в дальнейшем.

Эта тема является актуальной, так как любой полигон ТБО можно охарактеризовать с помощью многомерных векторов, в которых указаны данные об экологическом мониторинге и ресурсному потенциалу захороненных отходов. Следовательно, полигоны и свалки ТБО можно

изучать с помощью методов многомерного статистического анализа.

Цель данной бакалаврской работы – найти зависимость степени стабильности отходов на объектах захоронения ТБО от данных об экологическом мониторинге, который содержит значения химико-биологических показателей, морфологического и фракционного составов, с помощью методов многомерного статистического анализа.

Для достижения поставленной цели, необходимо решить следующие задачи:

- изучить теорию о твердых бытовых отходах и полигонах ТБО;
- изучить теорию о методах многомерного статистического анализа;
- проанализировать данные о химико-биологических показателях, фракционном и морфологическом составе проб отходов;
- провести кластерный анализ для классификации объектов по степени стабилизации отходов;
- определить зависимость фракционного состава отходов от возраста объекта и глубины отбора проб отходов;
- определить зависимость морфологического состава ТБО от возраста захоронения;
- построить регрессионную модель зависимости степени стабильности отходов от исследуемых показателей с помощью факторного анализа.

Объектами исследования выступают свалки и полигоны захоронения твердых бытовых отходов.

Предмет исследования – состояние объектов захоронения ТБО на основе данных о химико-биологических показателях, фракционном и морфологическом составе.

Данная работа состоит из введения, трех глав, заключения, списка литературы и приложений. В первой главе представлена теория по полигонам твердых бытовых отходов, их мониторингу. Во второй главе теоретические сведения о методах реализации поставленной цели (корреляционно-регрессионный анализ, кластерный анализ, дисперсионный анализ и факторный анализ). В третьей главе представлена практическая часть, в результате которой получена зависимость степени стабильности отходов на полигонах ТБО от химико-биологических показателей, данных морфологического и фракционного составов

ГЛАВА 1. ОСНОВНЫЕ СВЕДЕНИЯ О ПОЛИГОНАХ ТБО

1.1. Твердые бытовые отходы и полигоны ТБО

С каждым годом увеличивается площадь городов нашей страны, зачастую превращающихся в мегаполисы, неизменно растут потребности населения, что в свою очередь ведёт к увеличению числа предприятий, заводов. И, как следствие всего этого, растет количество потребляемых населением продуктов и, соответственно, количество выбрасываемого мусора. Число отходов с каждым годом возрастает на 3 %, одна четвертая часть которых приходится на твердые бытовые отходы.

Твердые бытовые отходы (ТБО, бытовой мусор) – это пищевые продукты, бытовые предметы, наибольшая часть продуктов потребления, которые потеряли свои качества, необходимые потребителю. Это те предметы, которые человек уже использовал, надобность в которых утрачена[1].

Основными источниками образования ТБО являются[2]:

- индивидуальные и многоквартирные жилые дома;
- здания общественного пользования;
- магазины, торговые залы и помещения;
- места, где проводятся концерты, зрелищные выступления, аттракционы, парки культуры и отдыха и тому подобное.

Все ТБО принято подразделять на:

- биологические отходы (пищевые отходы – еда);
- бытовой мусор (стекло, макулатура, пластик и др.)

Также бытовые отходы можно классифицировать по разным факторам. Например, по составу, времени использования, степени опасности.

Так ТБО по морфологическому составу принято подразделять:

- макулатура (бумага, картон);
- полимерные материалы (полиэтилен, пенопласт, пластмасса);
- металлы (черные и цветные);
- стекло;
- дерево;
- кожа;

- растительные отходы (листья, трава);
- пищевые отходы;
- резина (автомобильные покрышки);
- строительный мусор;
- опасные БО (растворители, ядохимикаты, лаки);
- текстиль;
- остаточные материалы (зола, ил, пыль);
- кости;
- неорганические отходы (камни, песок, керамика);
- прочие отходы.

Каждый год в России производится около 6,2 млрд тонн отходов, из которых 63 млн тонн составляют твердые бытовые отходы. Это, в среднем, составляет 445 кг на человека[2].
Состав ТБО:

Таблица 1.1

Состав ТБО в России

Вид ТБО	Доля вида отходов в общей массе ТБО, %
Макулатура	35
Пищевые отходы	41
Полимеры	3

Продолжение таблицы 1.1

Стекло	8
Металлы	4
Текстиль и другое	9

По длительности использования ТБО разделяют:

- кратковременные (еда, одежда, бумага и т.д.);
- товары длительного пользования (электроника, шины и пр.)

Классы опасности ТБО[3]:

- I класс. К нему относятся те бытовые отходы, которые могут нанести серьезный вред людям и окружающей среде. Природа после воздействия токсичных соединений мусора, принадлежащего данному классу, не может восстановиться. К данному классу опасности относятся материалы, содержащие ртуть (люминесцентные лампы, ртутные термометры и др.);
- II класс. Сюда также относятся опасные предметы, которые отрицательно влияют на природу, но после их воздействия она может в дальнейшем медленно восстановиться (от 30 лет и более), если человек вовремя сможет предпринять меры по удалению отрицательного воздействия отходов данного класса. К нему относятся машинные масла, аккумуляторы, элементы питания – батарейки и пр.
- III класс. Это материалы, способные причинить вред окружающей среде, но при этом природа способна восстановиться за промежуток времени в пределах одного десятилетия. К этому классу относится так называемый строительный мусор (краски, цемент, ацетон, металлосодержащая продукция).
- IV класс. Сюда относятся почти безвредные для человек и природы отходы. Если окружающая среда оказалась загрязненной мусором, относящемуся к данному классу, то она может полностью восстановиться за 3 года. Этот

класс составляют такие бытовые отходы как бумага, древесина, пластик и т.д.

- V класс. К данному классу относятся материалы, безвредные и для человека, и для окружающей его природы. Это пищевые отходы, керамика и пр.

В настоящее время существует большое количество несанкционированных свалок, на которые индивидуальные лица и фирмы выбрасывают свой мусор. Это чревато не только ухудшением экологического состояния окружающей среды, но и возникновением всевозможных инфекционных заболеваний, которые могут обернуться всевозможными эпидемиями и даже пандемией, заражением грунтовых вод, водоемов, почвы и воздуха, что в целом пагубно скажется на здоровье местных жителей и нации в целом. Неправильное обращение с отходами может привести к пожарам, а, как известно из многочисленных возгораний свалок, победить данное стихийное бедствие очень сложно и требует больших затрат материальных и человеческих ресурсов. Поэтому для хранения твердых бытовых отходов необходимо использовать такие специализированные сооружения и площадки, которые бы минимизировали негативное влияние ТБО на окружающий мир. В настоящее время самыми простыми и популярными подобными сооружениями являются полигоны ТБО.

Полигон ТБО представляет собой комплекс природоохранных сооружений, которые предназначены для складирования, изоляции и обезвреживания твердых бытовых отходов, где учитываются все процессы, которые происходят

с мусором во время хранения. Данные сооружения должны обеспечивать статическую устойчивость ТБО.

Для полигонов ТБО существует ряд требований к их устройству, которые прописаны в СНиП, СанПиН и прочих законах, касающихся охраны окружающей среды и градостроительства.

При выборе земельного участка для размещения полигона ТБО необходимо учитывать климатические, географические, гидрогеологические особенности местности, характеристики почвы. Запрещается размещать полигоны ТБО на территории первого и второго поясов ЗСО (зон санитарной охраны) водных и минеральных источников, во всех зонах охраны курортов, в местах массового отдыха населения, оздоровительных учреждений[4].

Санитарно-защитная зона между жилой застройкой и границами полигона должна быть не менее 500 м. Этот размер может изменяться при расчете газообразных выбросов в атмосферу.

На участке, на котором намечено размещение полигона ТБО, проводят соответствующие исследования и изыскания. Если на этом участке почва с большим содержанием глины или тяжелых суглинков, грунтовые воды расположены на глубине более 2 м, то это является лучшим основанием для размещения на нем полигона твердых бытовых отходов.

Прежде чем участок земли превратиться в полигон ТБО, его специальным образом подготавливают. Прежде всего его застилают специальным материалом. На многих полигонах гидроизоляционным материалом является ткань, сделанная с

применением бентонитовой глины. Этот материал предотвращает проникновению вредных веществ от мусора вместе с водой в почву, тем самым защищая ее от вредного воздействия ТБО.

Если влага не проникает в почву, то она остается сверху гидроизоляционного слоя. Для того, чтобы вода там не скапливалась, устанавливают специальную дренажную систему. Со дна полигона жидкость, при помощи насосов и специального комплекса труб с отверстиями, поднимается на поверхность.

Итак, настоящий полигон твердых бытовых отходов должен выглядеть так: на дне котлована выложена бентонитовая ткань, затем дренажная система, сверху нее слой гравия для того, чтобы мусор не забивал трубы, и только после этого выкладываются твердые бытовые отходы, причем тоже слоями, чередуя слои мусора и песка. Пересыпая ТБО песком, полигон защищается от пожаров в летний период.

На любом полигоне твердых бытовых отходов обязательно должны находиться:

- непосредственно сам участок складирования мусора, который должен быть оснащен соответствующим ограждением;
- территория, на которой находится цех по сортировке мусора;
- инженерные сооружения, которые в полной мере обеспечивают деятельность полигона (сюда относятся и линии электропередач, подъездные дороги);

- административно-хозяйственная зона, на которой должны располагаться производственно-бытовые здания для персонала, помещения для приема пищи, туалеты, гаражи и навесы для машин и оборудования;
- площадь, с расположенной на ней лабораторией для проведения быстрой экспертизы.

1.2. Экологический мониторинг. Основные химико-биологические показатели

Полигоны твердых бытовых отходов являются местом сосредоточения опасных химических элементов, которые могут негативно влиять на окружающую среду. В результате различных химических и биологических изменений (разложение) выделяются различные химические соединения. Наблюдается значительное выделение тепла и газа.

На территории захоронения ТБО и в пределах их воздействия на окружающую среду обязаны проводить мониторинг, в ходе которого изучается состояние окружающей среды. Данный экологический мониторинг позволяет выявить степень стабилизации полигона.

Стабильным (устойчивым) полигоном захоронения ТБО является полигон, в котором масса захороненных отходов находится в стабильном состоянии, т.е. все показатели загрязняющих веществ не превышают экологически приемлемый уровень и не оказывают отрицательного воздействия на окружающую среду и здоровье людей [5].

В ходе этого мониторинга изучаются различные химико-биологические показатели. Рассмотрим некоторые из них.

Дыхательная активность (RI4) – это биологический тест на респираторную активность. С помощью него оцениваются способности к биологическому разложению органического вещества в аэробных условиях. Он позволяет определить степень потребности отобранной пробы твердых бытовых отходов в кислороде для биологического преобразования органического вещества в оптимальных условиях. Определяет стабильность отходов.

GS21 – коэффициент производства газа, характерный показатель для производства свалочного газа. Экспериментально показано, что данный показатель хорошо коррелирует ($R^2 = 0,92$) с частотой дыхания (RI4), а также с общим содержанием органических веществ в пробе [6].

ХПК и БПК – одни из важных показателей уровня загрязнённости сточных вод частицами органического происхождения. ХПК – это показатель химического потребления кислорода, а БПК – показатель биохимического потребления кислорода.

ХПК – значение, характеризующее суммарное содержание в воде органики по объёму израсходованного на их полное окисление кислорода.

БПК – это показатель, оценивающий объём кислорода, который потребовался для аэробов на окись, разложение органических веществ. Вычисляется путём замера количества кислорода, использованного без доступа света за конкретное время.

ХПК показывает общий объём органики в стоках, а БПК оценивает загрязнение конкретного объема воды [7].

Зольность – массовая доля золы, процент содержания негорючего материала, который остается от пробы при сгорании. Почти все частицы органического происхождения горят, поэтому можно сделать вывод, что зольность равна разнице общего объёма пробы (100%) и процентного содержания в ней органики.

Аммонийный азот – это опасный загрязнитель сточных вод. Его высокое содержание в отходах говорит о процессах биодеструкции ТБО, которые легко разлагаются, с образованием жирных кислот, аминокислот, глицерина, полисахаров, аммиака [8].

Энтальпия – это такое термодинамическое свойство вещества, которое указывает количество энергии, содержащееся в его молекулярной структуре. Из физики известны два вида энергии вещества: внутренняя и кинетическая. Часть первой постоянно остается в самом веществе для поддержания структуры, а часть кинетической энергии вещества становится недоступна, если его температура приближается к температуре окружающей среды. Энтальпия показывает то количество энергии, которая доступна для преобразования в теплоту при постоянном давлении и температуре [9].

Эти параметры используются для оценки степени биохимической стабильности твердых бытовых отходов на полигоне. Энтальпия еще и показывает, сколько тепла и энергии мы можем получить из ТБО путём его сжигания.

В экологический мониторинг так же входят определения морфологического и фракционного состава ТБО. Фракционный состав несет в себе информацию о содержании в пробах отходов фракций различного размера. Отходы просеивают через сито с различными размерами ячеек и определяют процентное содержание различных фракций. Морфологический состав несет в себе информацию о содержании в массиве полигона того или иного вида ТБО. Он определяется после исследования фракционного состава путем ручной переборки фракций и изучения их состава.

1.3. Рекультивация полигонов ТБО

Несмотря на то, что в настоящее время развиваются технологии по вторичной переработке мусора, огромная часть отходов человеческой жизнедеятельности в конечном итоге оказывается на полигонах ТБО. Если верить цифрам, то в Российской Федерации переработке подвергается только 4% отходов, а 96% отправляются на свалки. Каждый полигон имеет свой срок работы. Если объемы мусора достигают своих предельных показателей, которые были оговорены при проекте полигона, или подошел к концу срок эксплуатации, то полигон ТБО закрывают, он прекращает функционировать. Далее по плану должна проходить его рекультивация, которая является закономерным финалом работы каждого полигона[4].

Проект рекультивации должен разрабатываться еще на этапе создания самого полигона. Она нужна для того, чтобы

участок, на котором когда-то располагался полигон ТБО, можно было бы использовать дальше (например, построить на его месте парк). Рекультивацию начинают после того, как весь мусор пришел в свое стабильное состояние.

Процесс рекультивации полигона можно условно разделить на два основных этапа:

- **технический.** Сюда относятся монтаж систем для сбора и обезвреживания свалочного газа и фильтрата, выравнивание рельефа и террасирование склонов (создание на склонах ступеней), если этого не сделать, на полигоне будут происходить обвалы. Входят сюда также и покрытие поверхностей слоем непроницаемых материалов;
- **биологический.** Это завоз и подготовка почвы, посадка растений. Задача данного этапа – это восстановление хозяйственной ценности земли.

Когда выполнены эти два этапа рекультивации полигона ТБО, то возникает вопрос: что делать дальше с территорией бывшего полигона? Каким образом использовать новую площадь и что на ней можно возвести?

Специалисты предупреждают, что строительство на бывших полигонах стоит вести максимально аккуратно. На полигонах нельзя строить капитальные сооружения (жилищные и промышленные строения), но можно строить некие зоны отдыха (ландшафтные парки, спортивные сооружения), т.е. такие вещи, которые не требуют фундамента, так как его заложить на полигоне практически невозможно.

Хорошим примером использования территории рекультивированного полигона является бывший полигон «Любарс». Он был построен в 1957 году в Берлине и проработал до 1982 года, когда было принято решение о его закрытии, рекультивации и превращении в парк. Это были первые в мире попытки из горы мусора сделать безопасную территорию. Сначала ничего не получалось: первые деревья и растения, которые высадили здесь, погибли. Затем, несколько раз система перестраивалась, совершенствовалась, и в конце концов на 5 млн тонн мусора расцвел настоящий зеленый сад. Сейчас это место отдыха граждан, где по выходным ходят различные концертные мероприятия.

На южной окраине Берлина есть парк «Мариенфельд». И сложно представить, что это место было когда-то городской свалкой с 4 миллионами кубометров мусора. В отличие от парка «Любарс», он более интерактивный. Здесь можно посмотреть на животных, есть несколько искусственных прудов, между которыми проложен экологический маршрут. Многие местные жители даже не догадываются, что этот парк находится на вершине бывшего полигона ТБО.

Подобные проекты реализуются и в России. Один из примеров – это первое в нашей стране поле для гольфа, которое расположено на бывшей московской свалке.

Всесезонный горнолыжный комплекс «Лисья гора» так же расположен на месте бывшего полигона ТБО, работавшего в 1980-х годах.

1.4. Польза и выгода от ТБО.

Технология складирования твердых бытовых отходов на полигонах известна в нашей стране уже давно. С каждым годом число производимого мусора населением растет, растут и площади складирования мусора. Данная технология обращения с мусором стала неэффективной, так как ученые обнаружили, что отходы можно использовать как источник получения различных полезных веществ, топлива, энергии. Поэтому простое складирование отходов на полигонах ТБО с экономической точки зрения является неправильным.

На сегодняшний день широкое распространение получила технология полигонного захоронения отходов с целью получения свалочного газа, который образуется из-за разложения мусора по слоям грунта. Этот газ содержит в себе до 60% метана, благодаря чему может быть использован в качестве топлива. При разложении одной тонны ТБО образуется не менее 100-200 кубометров свалочного газа, который не просто находится в теле полигона, но и который постоянно генерируется, благодаря естественным процессам, проходящим внутри свалки.

Прежде чем свалочный газ превратится в энергию, проводятся следующие этапы:

- проводится оценка существующего потенциала той или иной свалочной территории;
- оценка параметров того газа, который генерируется произвольно на полигоне;

- технологическое решение по сооружению группы газосборных скважин;
- технология конденсации и сбора газа;
- проводятся различные мероприятия по обработке газа. Свалочный газ проходит через газосушитель, нагреватель и фильтрацию активированным углем;
- газ поступает на газопоршневые установки, на которых вырабатывается электрическая энергия и попадает в сеть.

Австралийская компания «ADGEXLimited» разработала уникальный мобильный комплекс «greenBLAZE» по переработке любых углеродсодержащих отходов с получением на выходе высококачественного топлива, а также электрической и тепловой энергии. Компания утверждает, что они могут очистить территорию России от любых свалок. Данная система высокоэффективна и позволит снизить уровень воздействия на окружающую среду до минимума.

В специальную камеру помещается мусор. При помощи метода вакуумной деструкции (высокая температура и низкое давление) мусор переходит в газообразное состояние (синтез-газ), при этом не происходят выбросы опасных веществ в атмосферу. Из такого газообразного состояния мусора можно получить любой вид высокоуглеродного соединения – синтетическое моторное топливо (керосин, бензин, дизель).

Плюсы данной программы для экологии и общества заключаются в следующем:

- отсутствие выбросов вредных фуранов, диоксинов и иных опасных компонентов;

- высокий уровень безопасности установки;
- позволяет ликвидировать свалки и захоронения мусора;
- способствует сохранению экологического баланса окружающей среды;
- сохраняет биологические ресурсы Планеты.

RDF (RefuseDerivedFuel, альтернативное топливо) – это топливо, полученное из отходов. Для его производства используются следующие фракции мусора и отходов: резина, бумага, бытовые органические отходы, дерево, картон, кожа (кожзаменители), синтетические волокна, текстиль, пластик и т.д. Все то, что содержится в свалках кроме стекла, камней, металлов и хлорсодержащих материалов. Этот вид топлива часто используют в качестве дополнительного топлива на ТЭЦ, металлургических заводах, цементных заводах, котельных установках. RDF получают путем измельчения, сепарации и обезвреживания ТБО. В процессе производится горючая, высококалорийная фракция размером 20-25 мм. Сжигание такого вида топлива оказывает значительно меньшее воздействие на окружающую природу, чем простое сжигание ТБО на мусоросжигательных заводах [10].

Любой полигон ТБО можно охарактеризовать с помощью многомерных векторов, в которых указаны данные об экологическом мониторинге и ресурсному потенциалу захороненных отходов. Следовательно, полигоны и свалки ТБО можно изучать с помощью методов многомерного статистического анализа.

ГЛАВА 2. ПРИМЕНЯЕМЫЕ МЕТОДЫ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА

2.1. Корреляционно-регрессионный анализ

Корреляционно-регрессионный анализ – один из наиболее широко используемых методов многомерного статистического анализа, методов обработки исходных статистических данных.

Корреляционный анализ подразумевает под собой исследование взаимосвязи показателей исследуемого процесса. Самыми важными показателями в корреляционном анализе являются значения коэффициентов корреляции между парами признаков.

Коэффициенты парной корреляции – показывают степень линейной зависимости между различными парами факторов. При расчете коэффициентов парной корреляции учитывается и то, что взаимосвязь каждой пары исследуемых признаков находится под воздействием связи между собой и с признаками данной пары остальных признаков. Рассмотрим, как получается матрица парных корреляций исходя из матрицы исходных данных[11]:

- для начала исходную матрицу признаков необходимо стандартизировать, в результате чего получаем матрицу Z , в которой математические ожидания каждого признака равны 0, а дисперсии равны 1;

- затем находится произведение двух матриц: $Z'Z$ – транспонированной матрицы нормированных данных и самой матрицы Z ;
- искомая матрица парных корреляций находится по формуле: $R = \frac{1}{n} Z'Z$.

Помимо парных коэффициентов корреляции, корреляционный анализ данных подразумевает под собой поиск коэффициентов частных корреляций. Они, также как и коэффициенты парной корреляции, представляют собой линейную взаимосвязь пары признаков, но при этом связи всех других признаков с факторами данной пары не учитываются. Элементы матрицы коэффициентов частной корреляции можно получить с помощью матрицы парных корреляций:

$$r_{ij} = \frac{-A_{ij}}{\sqrt{A_{ii}A_{jj}}}, \quad (2.1)$$

где A_{kl} – алгебраические дополнения к соответствующим элементам матрицы парных корреляций.

Важным этапом корреляционно-регрессионного анализа перед построением регрессии является проверка того, как сильно исследуемый фактор зависит от других показателей в совокупности. Для этого используют коэффициент множественной корреляции и коэффициент детерминации.

Коэффициент множественной корреляции дает нам информацию о степени линейной зависимости объясняемой переменной от совокупности других (объясняющих) переменных. Данный показатель корреляционно-

регрессионного анализа рассчитывается с помощью матрицы парных корреляций по формуле:

$$R_0 = \sqrt{1 - \frac{|R|}{|R_j|}}, \quad (2.2)$$

где $|R|$ – определитель матрицы парных корреляций;

$|R_j|$ – минор к матрице парных корреляций. В матрице вычеркиваются строка и столбец, соответствующие характеристикам связи с j -м признаком.

Коэффициент детерминации равен квадрату коэффициенты множественной корреляции и показывает, на сколько процентов данная регрессионная модель близка к линейной. Или, другими словами, какова теснота линейной связи между откликом и другими признаками. Численная характеристика доли вариации отклика, объясненной вариацией всех признаков, включенных в модель.

Регрессионный анализ позволяет исследователю найти вид линейной зависимости исследуемого признака от действия других факторов. В нем исследуется линейная зависимость результирующей переменной (отклика, объясняемой переменной) y от нескольких объясняющих переменных (x_j), то есть регрессия вида [12]:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_m x_m + \varepsilon, \quad (2.3)$$

Так же в регрессионном анализе учитывается возможность наличия мультиколлинеарности среди объясняющих переменных. Мультиколлинеарностью называется явление, при котором наблюдается тесная связь (по данным парных корреляций) между объясняющими переменными [13].

Чтобы найти параметры (или коэффициенты) регрессионной модели b_k , используют метод наименьших квадратов (МНК). Согласно этому методу, минимизируется сумма квадратов отклонений исходных данных отклика от модельных значений \hat{y} [14]:

$$\sum \varepsilon^2 = \sum (y - \hat{y})^2 \rightarrow \min. \quad (2.4)$$

Вектор неизвестных параметров находится по формуле:

$$B = \hat{X}^{-1} Y \quad (2.5)$$

где X - матрица исходных данных, которая имеет вид:

$$X = \begin{pmatrix} \hat{x}_{11} & \hat{x}_{12} & \dots & \hat{x}_{1m} \\ \hat{x}_{21} & \hat{x}_{22} & \dots & \hat{x}_{2m} \\ \hat{x}_{\dots} & \dots & \dots & \dots \\ \hat{x}_{n1} & \hat{x}_{n2} & \dots & \hat{x}_{nm} \end{pmatrix}. \quad (2.6)$$

По данным формуле и матрице исходных параметров в ходе решения регрессионного уравнения находят все b_j , кроме b_0 , а затем b_0 вычисляется как разность: $b_0 = \bar{y} - \bar{X}' B^i$,

$$b_0 = \bar{y} - \bar{X}' B^i, \quad (2.7)$$

где \bar{y} - среднее значение отклика;

\bar{X}' - вектор средних значений признаков $\bar{X}' = (\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m)$;

B^* - неполный параметрический вектор, $B^{i*} = (b_1, b_2, \dots, b_m)$.

В найденной регрессионной модели (уравнении в натуральном масштабе) коэффициенты регрессии показывают, на сколько натуральных единиц изменится объясняемая переменная, если увеличить значение соответствующего признака на единицу.

Также, помимо модели в натуральном масштабе, можно построить регрессию, которая основывается на стандартизированной матрице исходных данных:

$$\hat{Z}_0 = \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_m Z_m. \quad (2.8)$$

Нахождение неизвестных параметров также основывается на МНК. Свободный член данной регрессии будет равен 0. Здесь коэффициенты регрессии показывают, на сколько средних квадратических отклонений изменится отклик при изменении соответствующего признака на одно среднее квадратическое отклонение.

От стандартизированных коэффициентов регрессии всегда можно осуществить переход к коэффициентам в натуральном масштабе: $b_j = \beta_j \frac{\sigma_0}{\sigma_j}$, где σ_0 – среднее квадратическое отклонение значений отклика Y ;

σ_j – среднее квадратическое отклонение значений соответствующего x_j .

После построения регрессии подразумевается ее проверка на надежность и качество решения. В ходе этих проверок исследуется качество всей построенной модели и качества отдельно взятого коэффициента модели.

Надежность регрессионной модели можно проверить непосредственно по значению коэффициента множественной корреляции и коэффициента детерминации. Рассмотрим значения, которые может принимать коэффициент детерминации [11]:

- 0,01-0,2 – связь отклика и признаков слабая, теоретически подтверждена недостаточно;
- 0,2-0,5 – связь средняя;
- 0,49-1,00 – связь достаточно сильная, использование регрессионной модели в анализе теоретически обосновано.

Статистический критерия Фишера отвечает за качество всей модели в целом.

$$F_n = \dots \quad (2.9)$$

Значения F -критерия F_n сравниваются с критическим (табличным) при заданном уровне значимости α (вероятность отвергнуть правильную гипотезу) и числе степеней свободы $\nu_1 = m$ и $\nu_2 = n - m - 1$. Надежность регрессионной модели подтверждается при $F_n > F_{табл.}$

С помощью t -критерия Стьюдента проверяется надежность отдельного коэффициента регрессии. Для некоторого j -го параметра регрессионной модели вычисляется наблюдаемое значение t -критерия по формуле:

$$t_{n(j)} = \frac{b_j}{\hat{S}_j}, \quad (2.10)$$

где \hat{S}_j – средняя ошибка j -го коэффициента регрессии $\hat{S}_j = \dots$, дисперсионная характеристика $\hat{S}_{оcм}^2 = \dots$ а c_{jj} – соответствующие (b_j) диагональные элементы матрицы \dots , при условии, что X – расширенная матрица, то есть в начало добавляется столбец, полностью состоящих из 1.

Наблюдаемые значения t -критерия также сравниваются с табличным при заданном α и числе степеней свободы $\nu = n$

- t - 1. Значимость коэффициента регрессии подтверждается, когда $t_n > t_{табл}$.

2.2. Кластерный анализ как метод группировки статистических данных

Кластерный анализ является еще одним методом обработки исходных данных. Это метод многомерного статистического анализа, позволяющий классифицировать исходные данные. Кластерный анализ позволяет рассматривать большой объем информации и классифицировать многомерные наблюдения, каждое из которых описывается набором исходных переменных.

Пусть имеется некоторое множество A объектов a_i , характеризующихся вектор-функцией $X_i = (x_{i1}, x_{i2}, \dots, x_{is})$, $i = 1, \dots, n$. Эти объекты характеризуются m признаками и представляют собой точки в m -мерном пространстве.

Целью кластерного анализа является образование групп схожих между собой объектов (кластеров) $A_k, k = 1, \dots, r$, где r - заданное число подмножеств. Для этих разбиений должны выполняться следующие свойства:

- $A_k \neq \emptyset$, это значит, что каждый кластер должен содержать по крайней мере 1 элемент;
- $A_k \cap A_l = \emptyset, k \neq l$, то есть каждый элемент должен принадлежать только одному кластеру;
- $\sum_{k=1}^r A_k = \bigcup_{k=1}^r A_k = A$, объединение всех кластеров образует исходное множество объектов.

Методы кластерного анализа позволяют решить следующие задачи [11]:

- проведение классификации объектов с учетом признаков, отражающих природу объектов;
- проверка выдвигаемых гипотез о наличии некоторой структуры в изучаемой совокупности объектов.

Каждый кластер имеет такие математические характеристики как центр, радиус, среднеквадратическое отклонение, размер кластера.

Центр кластера – это среднегеометрическое место расположения точек (наблюдаемых объектов) в пространстве.

Радиус кластера – это максимальное значение расстояний от каждого объекта до центра кластера. Объект относится к данному кластеру, если его расстояние до центра меньше радиуса этого кластера. [15]

Размер кластера может характеризоваться либо среднеквадратическим отклонением, либо радиусом кластера.

Методы кластерного анализа (методы объединения объектов в кластеры) можно разделить на иерархические и итеративные. Иерархические, в свою очередь, делятся на дивизимные и агломеративные

Оценка сходства объектов между собой тесно связана с абсолютным значением признака и степени его вариации в совокупности. Чтобы данные свойства не влияли на качество кластеризации, исходные значения показателей принято нормировать одним из способов нормировки [11]:

- $z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_i},$ (2.1 1)
- $z_{ij} = \frac{x_{ij}}{x_{max j}},$ (2.1 2)
- $z_{ij} = \frac{x_{ij}}{\bar{x}_j},$ (2.1 3)
- $z_{ij} = \frac{x_{ij}}{x_{min j}},$ (2.1 4)

В анализе кластеризации для количественной оценки сходства объектов в кластере вводится понятие метрики, с помощью которой определяется сходство или различие между классифицируемыми объектами. Используются различные меры расстояния между объектами [11]:

- евклидово расстояние: $d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2};$
- взвешенное евклидово расстояние: $d_{ij} = \sqrt{\sum_{k=1}^m w_k (x_{ik} - x_{jk})^2};$
- расстояние городских кварталов (манхэттенское расстояние): $d_{ij} = \sum_{k=1}^m |x_{ik} - x_{jk}|;$
- расстояние Минковского: $d_{ij} = \sqrt[p]{\sum_{k=1}^m |x_{ik} - x_{jk}|^p},$

где d_{ij} – расстояние между объектами i и j ;

x_{ik}, x_{jk} - значения переменной k у объектов i и j соответственно;

w_k – вес, приписываемый переменной k .

В случае зависимых признаков и их различной значимости при кластеризации объектов за меру близости объектов принимают расстояние Махаланобиса[16]:

$$d_{ij} = (X_i - X_j)^T S_i^{-1} (X_i - X_j),$$

где X_i и X_j – векторы значений переменных у объектов i и j соответственно;

S_i – общая ковариационная матрица.

Рассмотрим функцию меры различия элементов $\rho(X_i, X_j)$, которая удовлетворяет условиям:

- $\rho(X_i, X_j) \geq 0, \forall i, j$ – неотрицательность функции различия;
- $\rho(X_i, X_j) = 0, \forall i = j$ – эквивалентность (тождественность) элементов;
- $\rho(X_i, X_j) = \rho(X_j, X_i)$ – симметричность меры различия;
- $\rho(X_i, X_j) \leq \rho(X_i, X_k) + \rho(X_k, X_j)$ – неравенство треугольника.

Данная функция различия – это обобщение функции евклидова расстояния:

$$\rho(X_i, X_j) = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2}, \quad (2.15)$$

Так же в кластерном анализе рассматривается и функция сходства:

$$s_{ij} = \frac{1}{1 + \rho(X_i, X_j)}, \quad (2.16)$$

2.2.1. Иерархические методы кластерного анализа

Как было сказано выше, иерархические методы кластерного анализа делятся на дивизимные и агломеративные процедуры. В дивизимных (или разделяющих) методах сначала все объекты объединены в один общий кластер, а на последнем шаге каждый объект – это отдельный кластер. В агломеративных процедурах все

наоборот, сначала каждый объект представляет собой отдельный кластер, а на последнем шаге все объекты собраны в один кластер. Иначе, дивизимная процедура – это разбиение совокупности объектов на кластеры, а агломеративная процедура – объединение объектов в кластер [17].

Агломеративный иерархический кластерный анализ можно описать следующим алгоритмом:

- самым первым шагом является стандартизация данных одним из рассмотренных выше способов;
- следующим действием является расчет матрицы расстояний или мер сходства между кластерами;
- выбирается пара объектов, которые ближе всего расположены друг к другу, и объединяются в один кластер под номером, меньшим из номеров объединяемых кластеров;
- второй и третий шаги повторяются до того, как все объекты не будут объединены в один единственный кластер.

Дивизимный иерархический кластерный анализ, в свою очередь, может быть представлен следующим алгоритмом:

- самым первым шагом является стандартизация данных одним из рассмотренных выше способов;
- следующим действием является расчет матрицы расстояний или мер сходства между кластерами;
- по матрице расстояний находятся два самых дальних элемента в каждом кластере, и тот кластер, где

расстояние между дальними объектами максимально, разбивается на 2 новых кластера. Оставшиеся элементы присваиваются тому кластеру, к которому ближе всего расположены;

- второй и третий шаги повторяются до того момента, когда все объекты будут разбиты на отдельные кластеры [11].

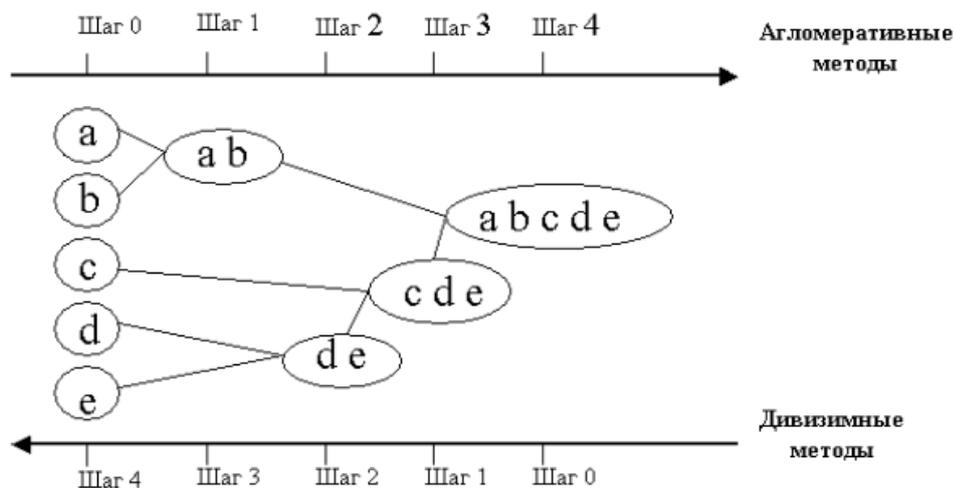


Рисунок 2.1 Схема итеративных методов

Иерархические методы используются при небольших объёмах выборки. Их преимуществом является наглядность кластеризации, так как данные алгоритмы связаны с построением дендрограмм (результат иерархического кластерного анализа).

Дендрограмма – это древовидная диаграмма, на которой изображаются этапы объединения данных в кластеры (или разбиения общего кластера на составляющие) [18].

Иерархические агломеративные методы кластерного анализа являются самыми популярными методами

кластеризации статистических данных. Между собой эти методы отличаются условиями объединения кластеров в один. Рассмотрим некоторые из этих методов.

Метод одиночной связи подразумевает под собой то, что кластер объединится с другим при условии, что он будет иметь максимальное сходство хотя бы с одним элементом другого кластера, к которому он собирается присоединиться.

При методе полных связей включение нового объекта к кластеру происходит только тогда, когда расстояние между ним и другими объектами исследуемого кластера не больше некоторого заданного уровня [18].

При методе средней связи, для выяснения того, будет ли включен новый объект в кластер, вычисляется среднее значение меры сходства, которое в последствии сравнивается с заданным пороговым значением.

При рассмотрении вопроса об объединении двух кластеров, то мерой служит расстояние между центрами исследуемых кластеров.

В методе Уорда для кластеров определяются средние значения каждого признака и рассчитывается сумма квадратов отклонений:

$$V_k = \sum_{i=1}^{n_k} \sum_{j=1}^m (x_{ij} - x'_{jk})^2, \quad (2.17)$$

где k – номер кластера;

i, j – номера объекта и признака соответственно;

m – количество признаков, характеризующих каждый объект;

n_k – количество объектов в кластере.

В последующих шагах объединяются те кластеры, которые дают наименьшее приращение величины V_k [11].

2.2.2. Метод k -средних как итеративный метод кластерного анализа

Наряду с иерархическими методами кластерного анализа применяют итеративные методы, к которым относится метод k -средних. Сущность таких методов кластеризации заключается в том, что процесс кластеризации начинается с задания некоторых начальных условий, в случае метода k -средних этим начальным условием является количество образуемых кластеров. Само название метода k -средних было предложено Дж. Мак-Куином в 1967 году. Данный метод не требует вычисления и хранения матрицы расстояний или сходств между объектами.

Данный метод так же можно представить в виде определенного алгоритма:

- все начинается с выбора необходимого числа кластеров, на которое необходимо разбить нашу совокупность исходных данных, задания первоначального разбиения и вычисления центров тяжести кластеров;
- далее происходит определение новых составов кластеров в соответствии с выбранной метрикой;
- следующим шагом является перерасчет центров тяжести кластеров;

- второй и третий шаги повторяются до тех пор, пока последующая итерация не даст аналогичный состав кластеров.

Описанный выше алгоритм называется алгоритмом Болла Холла. Метод k -средних по версии Мак-Куина отличается тем, что каждый раз, когда объект переходит из одного кластера в другой, происходит перерасчет эталонов (центров кластеров).

Рассмотрим следующую задачу и разберем ее по алгоритму Мак-Куина:

Пусть мы имеем n наблюдений. Каждый объект из этих наблюдений описывается m признаками. Задача состоит в том, что необходимо исходную совокупность разбить на k кластеров.

Применим алгоритм Мак-Куина:

- на начальном этапе из общего облака объектов выбираются случайным образом, или по усмотрению исследователя, k точек, которые будут являться центрами (эталоны) кластеров. Каждому центру присваивается свой номер;
- далее из оставшейся совокупности из $(n-k)$ точек выбирается любая точка и для того, чтобы определить к какому кластеру ее отнести, используется одна из метрик (например, евклидово расстояние). Исследуемая точка будет отнесена к тому кластеру, до эталона которого от этой точки расстояние будет наименьшим. Далее в этом кластере происходит перерасчет эталона с учетом нового объекта кластера;

- описанная выше процедура повторяется $(n-k)$ раз. В итоге мы получим, что все наши объекты будут разделены на k кластеров. Для того чтобы добиться устойчивости разбиения по тому же правилу, все точки X_1, X_2, \dots, X_n опять подсоединяются к полученным кластерам. Новое разбиение сравнивают с предыдущим. Если они совпадают, то работа алгоритма завершается. В противном случае цикл повторяется.

В обоих случаях итеративный алгоритм этого метода минимизирует дисперсию внутри каждого кластера, хотя в явном виде такой критерий оптимизации не используется[11].

2.3. Дисперсионный анализ как проверка качества кластеризации

Дисперсионный анализ – это один из методов многомерного статистического анализа, который применяется для анализа результатов наблюдений, зависящих от различных одновременно действующих факторов, цель которого состоит в выборе наиболее значимых, предпочтительных факторов и оценки их влияния на исследуемый процесс. Дисперсионный анализ позволяет ответить на вопрос о том, изменится ли дисперсия результатов эксперимента при изменении уровней изучаемого фактора (уровнями фактора называются фиксированные значения данного фактора). Если дисперсии при разных уровнях фактора будут значимо отличаться друг

от друга, то можно сделать вывод о том, что исследуемый фактор значимо влияет на среднее значение наблюдаемой переменной [19].

В дисперсионном анализе проверяется нулевая гипотеза о равенстве средних значений по каждому уровню исследуемого фактора:

$$H_0: m_1 = m_2 = \dots = m_p, \quad (2.18)$$

Альтернативной гипотезой, соответственно, будет являться предположение о нарушении хотя бы одного из этих равенств.

Рассмотрим следующую задачу: пусть на случайную величину X действует некоторый фактор F , который имеет p уровней, а количество наблюдений на каждом уровне одинаково и равно q . \bar{x} – генеральное среднее значение всех наблюдений. Необходимо узнать, является ли исследуемый фактор значимым и как сильно он влияет на исследуемый процесс.

Для решения этой задачи с помощью дисперсионного анализа необходимо найти следующие значения:

$SS_{total} = \sum_{j=1}^p \sum_{i=1}^q (x_{ij} - \bar{x})^2$ – общая сумма квадратов отклонений наблюдаемых значений от общего среднего;

$SS_{BG} = q \sum_{i=1}^p (\bar{x}_i - \bar{x})^2$ – факторная (или межгрупповая, *between-group*) сумма квадратов отклонений групповых средних от общего среднего. Данное значение характеризует рассеяние между группами;

$SS_{WG} = \sum_{j=1}^p \sum_{i=1}^q (x_{ij} - \bar{x}_j)^2$ – остаточная (или внутригрупповая, *within-group*) сумма квадратов отклонений наблюдаемых значений группы от своего группового среднего. Этот показатель характеризует рассеяние внутри групп.

Для вышеуказанных показателей свойственно равенство:

$$SS_{total} = SS_{BG} + SS_{WG}, \quad (2.19)$$

С помощью этих значений можно получить значения общей, факторной и остаточной дисперсий:

- $MS_{total} = \frac{SS_{total}}{N-1},$ (2.20)
- $MS_{BG} = \frac{SS_{BG}}{p-1},$ (2.21)
- $MS_{WG} = \frac{SS_{WG}}{N-p},$ (2.22)

Статистическим критерием проверки нулевой гипотезы является статистика Фишера (или *F*-критерий). Значение этой статистики вычисляется по формуле:

$$F = \frac{MS_{BG}}{MS_{WG}}, \quad (2.23)$$

эта статистика сравнивается с критическим значением критерия Фишера на желаемом уровне значимости и со степенями свободы, соответствующими факторной и остаточной дисперсиям.

Нулевая гипотеза принимается, если $F < F_{крит.}$

Если нулевая гипотеза справедлива, то факторная и остаточная дисперсии являются несмещенными оценками дисперсии общей совокупности. Следовательно, незначимо

отличаются друг от друга. Значит исследуемый фактор не влияет на изменчивость рассматриваемого признака.

2.4. Факторный анализ

Перед исследователем часто стоит задача изучения многопараметрических объектов и постоянно возникает вопрос: можно ли сократить размерность этих изучаемых объектов, отбросить часть параметров или заменить их числом функций, меньшим исходному числу размерности? Эти вопросы и ряд других позволяет решать факторный анализ. [20]

Факторный анализ – это один из методов многомерного статистического анализа, предназначенный для изучения взаимосвязей между значениями переменных. Он исследует внутреннюю структуру ковариационных и корреляционных матриц.

Первоначально, целью факторного анализа было построение математической модели способностей и поведения человека. Самые известные психологические теории принадлежат Ч. Спирмэну, С. Барту, Т. Келли, Л. Тэрстоуну, К. Холзингеру и Г. Томпсону [21]. Считается, что поворотным моментом для начала теории факторного анализа является 1904 год, когда вышла статья Чарльза Спирмэна, в которой он обсуждал составляющие интеллекта.

Факторный анализ позволяет решать большое количество задач. Самыми основными и важными являются следующие задачи[22]:

- сокращение числа переменных;
- построение новых обобщенных показателей;
- наглядное представление многомерных наблюдений (проецирование данных);
- преодоление мультиколлинеарности переменных в регрессионном анализе.

Рассмотрим общую модель факторного анализа.

Пусть у нас имеется n -мерное пространство признаков. И из этого пространства получается k факторов, которые сохраняют в себе всю важную информацию, которую несли исходные параметры.

В результате этого преобразования получается следующая модель:

$$X_i = \sum_{j=1}^k a_{ij} F_j + U_i, i=1, \dots, n. \quad (2.24)$$

Получается, что наши исходные параметры раскладываются на два значения:

- первое слагаемое – общности. Это та информация из исходных параметров, которая перенеслась в новое факторное пространство;
- второе слагаемое – то, что осталось в стороне, помехи. Это то, что наши новые факторы не смогли объяснить.

a_{ij} – это факторные нагрузки. Они совпадают с коэффициентами корреляции между построенными факторами и исходными признаками.

Дисперсия каждого исходного признака раскладывается на две составляющие:

$$S_{X_i}^2 = H_i^2 + S_{U_i}^2, \tag{2.25}$$

- общность – это та часть дисперсии, которая моделью объяснилась;
- специфичность – потерянная при построении модели часть исходной дисперсии признака.

$$H_i^2 = \sum_{j=1}^k a_{ij}^2, \tag{2.26}$$

Самый распространенный метод факторного анализа – это метод главных компонент. Он позволяет уменьшить число переменных, выбрав самые изменчивые из них [23].

Новые переменные:

- изначально все данные стандартизируются, чтобы их математические ожидания равнялись 0, а дисперсия 1;
- первая главная компонента устроена таким образом, чтобы вдоль нее изменчивость показателей была максимальной;
- вторая главная компонента перпендикулярна первой, то есть не коррелирует с ней, и имеет максимально возможную в этом случае дисперсию признаков;

- третья главная компонента должна не коррелировать с первой и второй и иметь максимально возможную дисперсию признаков и так далее.

Вращение факторов – это повышение контрастности решения и интерпретируемости модели. После реализации метода главных компонент видно следующее ограничение: объем объясняемой дисперсии сильно снижается с каждым следующим фактором.

Есть разные способы вращения главных компонент, но если нас интересуют контрастные решения, то часто используют вращение под названием Varimax. Это ортогональный метод вращения, который минимизирует число переменных с высокими нагрузками на каждый фактор. Данный метод позволяет упростить интерпретацию факторов. В результате этого вращения мы получаем, что объясненная факторами дисперсия распределяется более равномерно между построенными осями (главными компонентами). В результате вращения факторов меняется и матрица факторных нагрузок.

Существуют разные способы проверки качества построенной факторной модели.

К формальным способам проверки относятся тест КМО (Кайзера-Мейера-Олкина), который является критерием адекватности выборки, и критерий сферичности Бартлетта.

Критерий адекватности выборки КМО позволяет провести проверку того, насколько корреляцию между переменными можно объяснить другими (новыми) факторами. Он не зависит от объема выборки и для него

отсутствуют статистические таблицы значимости [24].
Статистика этого критерия вычисляется следующим образом:

$$MSA = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} q_{ij}^2}, \quad (2.26)$$

где r_{ij}^2 – квадраты элементов матрицы коэффициентов корреляции (кроме главной диагонали), q_{ij}^2 – квадраты соответствующих частных коэффициентов корреляции.

Данный критерий сравнивает коэффициенты корреляции со значениями частных коэффициентов корреляции. Небольшие значения данного критерия говорят о том, что корреляции нельзя объяснить другими переменными и факторный анализ использовать некорректно [25].

Таблица 2.1

Мера выборочной адекватности КМО

Значение статистики КМО	Оценка адекватности применения факторного анализа
Более 0,9	Безусловная адекватность
Более 0,8	Высокая адекватность
Более 0,7	Приемлемая адекватность
Более 0,6	Удовлетворительная адекватность
Более 0,5	Низкая адекватность
Менее 0,5	Факторный анализ неприменим к выборке

Тест Бартлетта также проверяет возможность построения факторной модели. Этот тест проверяет гипотезу о диагональности корреляционной матрицы, то есть то, что каждый признак связан только с самим собой. Если матрица корреляции имеет диагональный вид, то построение факторной модели не имеет смысла. Статистика данного критерия имеет вид:

$$\chi^2 = - \left[n - \frac{1}{6}(2m+5) \right] * \ln|R|,$$

где n – объем выборки, m – число переменных, $|R|$ – определитель корреляционной матрицы.

- Если корреляция между признаками незначима, то определитель матрицы будет близок к 1, что соответствует низким значениям статистики Бартлетта;
- если же корреляция сильная, то определитель матрицы корреляций будет близок к 0, что соответствует высоким значениям критерия.

Данная статистика сравнивается с критическим значением распределения χ^2 . Нулевая гипотеза о диагональной матрице принимается, если статистика получилась ниже, чем критическое значение. Следовательно, факторный анализ использовать нецелесообразно.

ГЛАВА 3. АНАЛИЗ СТАБИЛЬНОСТИ ПОЛИГОНОВ ТБО

3.1. Исходные данные

В качестве исходных данных были взяты результаты экологического мониторинга по важным химико-биологическим показателям стабильности ТБО. Для проведения комплексных исследований по оценке состояния объектов захоронения отходов был произведен отбор проб на 3 объектах захоронения ТБО Пермского края, 1 объекте захоронения ТБО Свердловской области (полигон захоронения ТБО п. Верхняя Пышма) и 1 объекте захоронения ТБО Московской области (полигон ТБО «Дубна Левобережная»). Исследуемые объекты характеризуются разным сроком складирования мусора. Со всех объектов захоронения были взяты пробы отходов посредством ковшового экскаватора по всей глубине массива полигона с интервалом 0,5-1 м. Сроки захоронения были определены при сортировке мусора из проб по информации на упаковках, этикетках. Если определить возраст по информации на мусоре было невозможно (что характерно для старых полигонов), то возраст отходов принимался, опираясь на даты начала, окончания эксплуатации и рекультивации объекта [8].

- На городской свалке ТБО г. Кунгур было отобрано 12 образцов с 4 участков. 8 проб из скважин № 1 и 2 были взяты с места, где осуществляется складирование мусора (возраст отходов 0-5 лет), 4 пробы из скважин № 3 и 4 взяты с участка, который был рекультивирован в

2007 г. засыпкой песчано-глинистого грунта (возраст отходов 30-50 лет). Эксплуатируется с 1965 г.

- На свалке ТБО г. Краснокамск было отобрано 12 образцов ТБО. Возраст захороненных отходов 10-50 лет. Период эксплуатации свалки с 1963 по 2008 г.
- На свалке ТБО «Голый Мыс» было отобрано 7 образцов. Данная свалка не эксплуатируется с 1982 г., следовательно, возраст захороненных отходов составляет более 40 лет. Свалка «Голый Мыс» рекультивирована, покрыта песчано-глинистым грунтом и большая ее поверхность покрыта зеленой растительностью.
- На полигоне ТБО «Верхняя Пышма» было взято 14 проб отходов. Данный полигон эксплуатируется с 1970 года, возраст отходов на нем составляет 1-30 лет.
- На полигоне ТБО «Дубна Левобережная» было отобрано 6 образцов. Этот полигон эксплуатировался с 1965 по 2014 г. Возраст отходов на нем составляет 4-50 лет.

Исследуем таблицу, содержащую данные о 51 пробе, взятой с 5 свалок и полигонов ТБО. Данная таблица содержит 10 признаков, информация о которых была получена в ходе экологического мониторинга (см. Приложение А). Дальнейшие исследования проводились с помощью пакета «STATISTICA».

Построим корреляционную матрицу (см. Приложение Б).

По данным корреляционной матрицы видно, что многие признаки сильно коррелируют друг с другом. Чтобы в дальнейшем качественно провести кластерный анализ,

необходимо избавиться от явления мультиколлинеарности. Для этого необходимо из исследования исключить сильно коррелируемые факторы на основе корреляционной матрицы и теоретических сведений об изучаемых факторах.

Между процентом органического вещества и процентом зольности в пробе существует линейная связь ($r = -1$), так как почти все органические вещества полностью сгорают, а зольность есть несгорающий остаток, получившийся после процесса горения, то сумма процентов зольности и органического вещества дает нам 100%. Следовательно, можно оставить только один параметр (например, процент органического вещества).

Параметры ХПК и БПК отвечают за степень загрязненности сточных вод, следовательно, можно оставить только один из них (например, ХПК)

Показатель GS21 имеет тесную связь с тестом RI4. Так как тест на респираторную активность (RI4) является наиболее предпочтительным индикатором оценки стабилизации отходов, чем тест на газообразование, то показатель GS21 можно убрать.

Таким образом, после исследования корреляционной матрицы, осталась таблица с размерами 51*7, в которой отображены информации по таким параметрам как относительная глубина отборы проб, возраст отходов в пробе, тест на дыхательную активность RI4, процент содержания органического вещества в пробе, ХПК, аммонийный азот и энтальпия.

3.2. Проведение кластерного анализа с целью определения стабильных объектов захоронения.

Целью проведения кластерного анализа является разбиение проб отходов на группы, каждая из которых соответствует определенному уровню стабилизации отходов.

Так как все алгоритмы кластерного анализа требуют оценок расстояний между кластерами или объектами, то при вычислении расстояний необходимо задать масштаб измерений. Каждый исследуемый показатель имеет свои единицы измерения, типы шкал, следовательно, изучаемые данные необходимо стандартизировать. Данная процедура в программе выполняется по формуле (2.11). После стандартизации все данные имеют среднее значение 0 и стандартное отклонение 1.

Определим, образуют ли исследуемые объекты «естественные» кластеры, которые могут быть осмыслены. С этой целью используются агломеративные иерархические методы кластерного анализа. Построим вертикальную дендрограмму, используя метод полной связи и манхэттенское расстояние:

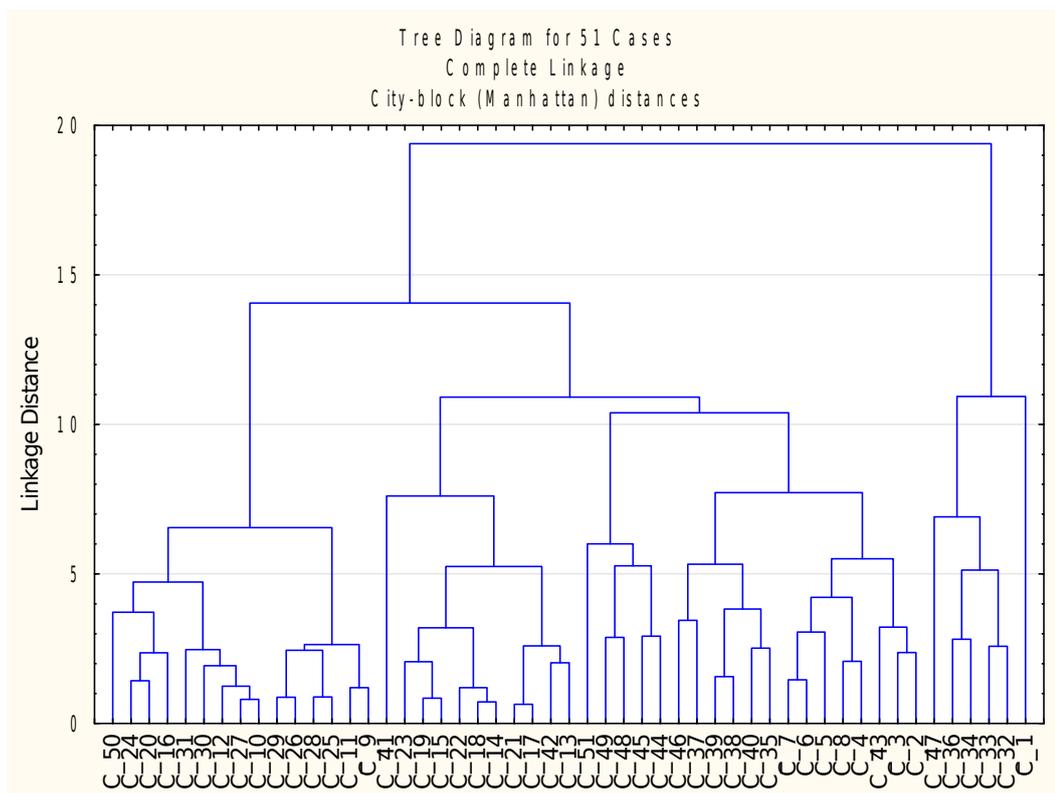


Рисунок 3.1 Дендрограмма

Исследуя дендрограмму, можно условно наши данные объединить в 5 кластеров.

Преимущество иерархических методов заключается в их наглядности. Метод итеративной кластеризации k -средних позволяет разбить данные на необходимое исследователю число кластеров. Его преимущество в том, что он позволяет исследовать кластеры, на которые разбились данные: проводить дисперсионный анализ, изучать близость кластеров, строить график средних значений параметров по каждому кластеру.

Проведем данный метод для различных значений K (пусть $K = 3, 4, 5$).

1) $K = 3$.

Таблица 3.1

Евклидовы расстояния между 3 кластерами

Cluster Number	Euclidean Distances between Clusters (TBO)		
	No. 1	No. 2	No. 3
No. 1	0,00000	0,84975	0,80550
No. 2	0,92182	0,00000	2,90162
No. 3	0,89749	1,70341	0,00000

В этой таблице показаны евклидовы расстояния между кластерами. Видно, что кластеры 1 и 3 ближе всего расположены друг к другу.

Таблица 3.2

Дисперсионный анализ разбиения на 3 кластера

Variable	Analysis of Variance (TBO)					
	Between SS	df	Within SS	df	F	signif. p
отн глуб отбора пробы	6,06536	2	43,93464	48	3,3133	0,044884
возраст	36,77121	2	13,22878	48	66,7113	0,000000
RI	28,90721	2	21,09278	48	32,8916	0,000000
% орг вещества	43,03885	2	6,96114	48	148,3863	0,000000
хпк	26,20653	2	23,79347	48	26,4340	0,000000
аммонийный азот	24,58170	2	25,41830	48	23,2101	0,000000
энгальпия	21,49666	2	28,50334	48	18,1003	0,000000

В данной таблице приведены результаты дисперсионного анализа. Параметры F и p характеризуют вклад признака в разделение объектов на группы. Лучшей кластеризации соответствуют большие значения F-критерия (проверка гипотезы о неравенстве дисперсий) и меньшие

значения p . Видно, что больше всего на разбиение данных по кластерам повлиял процент содержания органического вещества в пробе.

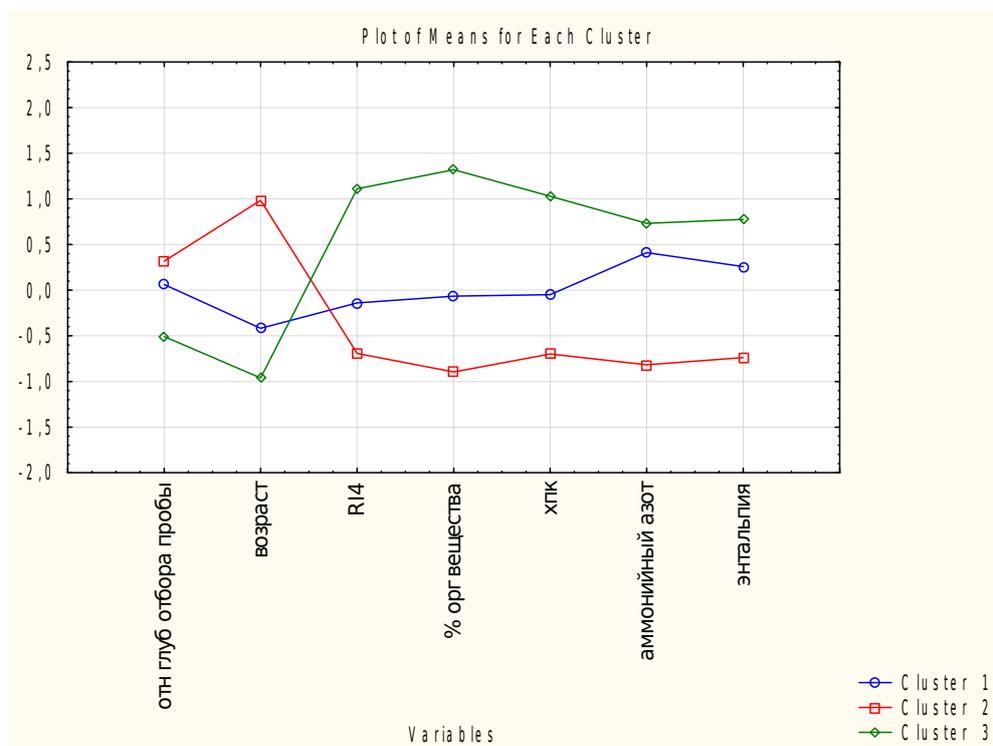


Рисунок 3.2 График средних значений показателей по 3 кластерам

Каждая ломаная линия на этом графике соответствует одному из кластеров. Каждое деление горизонтальной оси графика соответствует одной из переменных, включенных в анализ. Вертикальная ось соответствует средним значениям переменных для объектов, входящих в каждый из кластеров.

Обсуждение результатов разбиения данных (см. Приложение В):

- кластер №1 содержит информацию о 15 пробах отходов. Этот кластер характерен тем, что в нем все показатели приняли средние значения. К этому кластеру относятся части проб отходов, взятые с полигона ТБО «Верхняя Пышма», «Дубна Левобережная»;
- к кластеру № 2 относятся 21 проба, взятая с самых старых полигонов. На этих объектах все необходимые химико-биологические показатели пришли в стабильное состояние с окружающей средой. Сюда относятся такие объекты захоронения отходов как неэксплуатируемая часть городской свалки ТБО г. Кунгур, свалка ТБО г. Краснокамск и закрытая свалка «Голый Мыс»;
- кластер № 3 характерен тем, что пробы, относящиеся к нему, взяты с самых молодых полигонов. Эти полигоны имеют самые высокие значения химико-биологических показателей. Сюда относятся 15 проб отходов. К третьему кластеру относятся половина проб, взятых с эксплуатируемой части городской свалки ТБО г. Кунгур полигона ТБО «Верхняя Пышма» и «Дубна Левобережная».

2) Проведем кластеризацию данных при $K = 4$.

Таблица 3.3

Евклидовы расстояния между 4 кластерами

Cluster Number	Euclidean Distances between Clusters (TBO)			
	Distances below diagonal			
	No. 1	No. 2	No. 3	No. 4
No. 1	0,00000	1,18339	1,19261	0,62851
No. 2	1,08784	0,00000	0,87144	2,92117
No. 3	1,09206	0,93351	0,00000	3,20770
No. 4	0,79278	1,70914	1,79100	0,00000

Здесь дальше всего друг от друга расположены кластеры 3 и 4.

Таблица 3.4

Дисперсионный анализ разбиения на 4 кластера

Variable	Analysis of Variance (TBO)					
	Between S.S.	df	Within S.S.	df	F	signif. p
отн. глуб. отбора пробы	6,8235	3	43,1764	47	2,4759	0,07292
возраст	41,8512	3	8,1487	47	80,4621	0,00000
R.И	31,3055	3	18,6944	47	26,2351	0,00000
% орг. вещества	40,1737	3	9,8262	47	64,0519	0,00000
хПК	24,7327	3	25,2672	47	15,3352	0,00000
аммонийный азот	23,4388	3	26,5611	47	13,8249	0,00000
энтальпия	41,3942	3	8,6057	47	75,3577	0,00000

По данным дисперсионного анализа видно, что показатель относительной глубины отбора проб не повлиял на разбиение данных по кластерам.

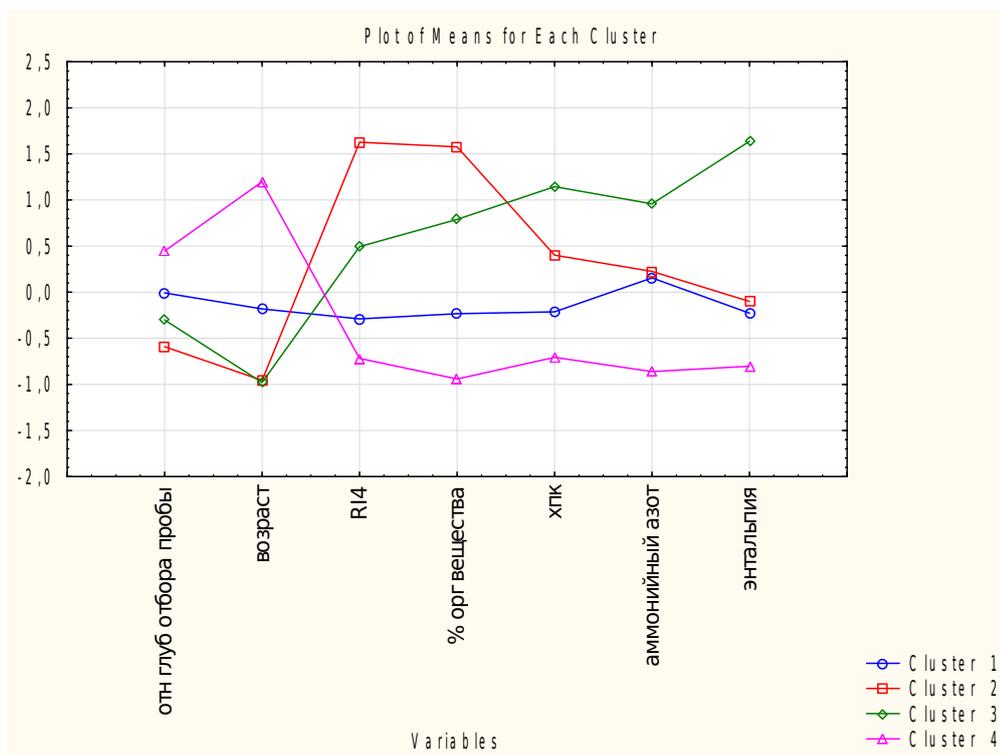


Рисунок 3.3 График средних значений показателей по 4 кластерам

Обсуждение результатов разбиения данных (см. Приложение Г):

- кластер №1 содержит информацию о 16 пробах отходов. Этот кластер характерен тем, что в нем все показатели приняли средние значения. К этому кластеру относятся части проб отходов, взятые с полигона ТБО «Верхняя Пышма», «Дубна Левобережная» и свалки ТБО г. Краснокамск;
- к кластеру № 2 относятся 7 проб отходов, взятые с самых молодых полигонов, отходы на которых обладают самыми высокими показателями теста на респираторную активность и процента органических

- веществ. К данному кластеру относится эксплуатируемая часть городской свалки ТБО г. Кунгур;
- кластер № 3 содержит 11 наблюдений и характерен для самых молодых полигонов ТБО. Пробы отходов взяты с полигонов с высокими показателями ХПК, аммонийного азота и энтальпии. Сюда относятся большая часть проб, взятых с полигона ТБО «Верхняя Пышма»;
 - кластер № 4 содержит 17 наблюдения и удовлетворяет стабильным объектам захоронения твердых бытовых отходов. Пробы отходов из этого кластера взяты со старых полигонов, все химико-биологические показатели пришли в стабильное состояние с окружающей средой. К этому кластеру относятся неэксплуатируемая часть городской свалки ТБО г. Кунгур, часть проб со свалки ТБО г. Краснокамск, свалка «Голый Мыс».
- 3) Проведем кластеризацию данных при $K = 5$.

Таблица 3.5

Евклидовы расстояния между 5 кластерами

Euclidean Distances between Clusters (TBOн)					
Cluster Number	Distances below diagonal				
	Squared distances above diagonal				
	No. 1	No. 2	No. 3	No. 4	No. 5
No. 1	0,00000	1,36495	0,72907	2,69310	2,16122
No. 2	1,16831	0,00000	1,56446	0,46704	0,98124
No. 3	0,85327	1,25078	0,00000	3,06896	2,20121
No. 4	1,64106	0,68340	1,75184	0,00000	0,65407
No. 5	1,47011	0,99057	1,48364	0,80875	0,00000

Дальше всего друг от друга расположены кластеры под номерами 3 и 4.

Таблица 3.6

Дисперсионный анализ разбиения на 5 кластера

Variable	Analysis of Variance (TBOн)					
	Between SS	df	Within SS	df	F	signif. p
Отн. глуб. отбора пробы	22,7290	4	27,2709	4	9,5846	0,00001
возраст	42,6682	4	7,3317	4	66,9255	0,00000
RI	34,2616	4	15,7383	4	25,0349	0,00000
% орг. вещества	37,4318	4	12,5681	4	34,2504	0,00000
элк	24,6428	4	25,3571	4	11,1760	0,00002
аммонийный азот	33,0662	4	16,9338	4	22,4557	0,00000
энталгия	37,8676	4	12,1323	4	35,8938	0,00000

По данным дисперсионного анализа можно сделать вывод о том, что сильнее всего на разбиение данных по кластерам повлияла переменная, отвечающая за возраст объекта.

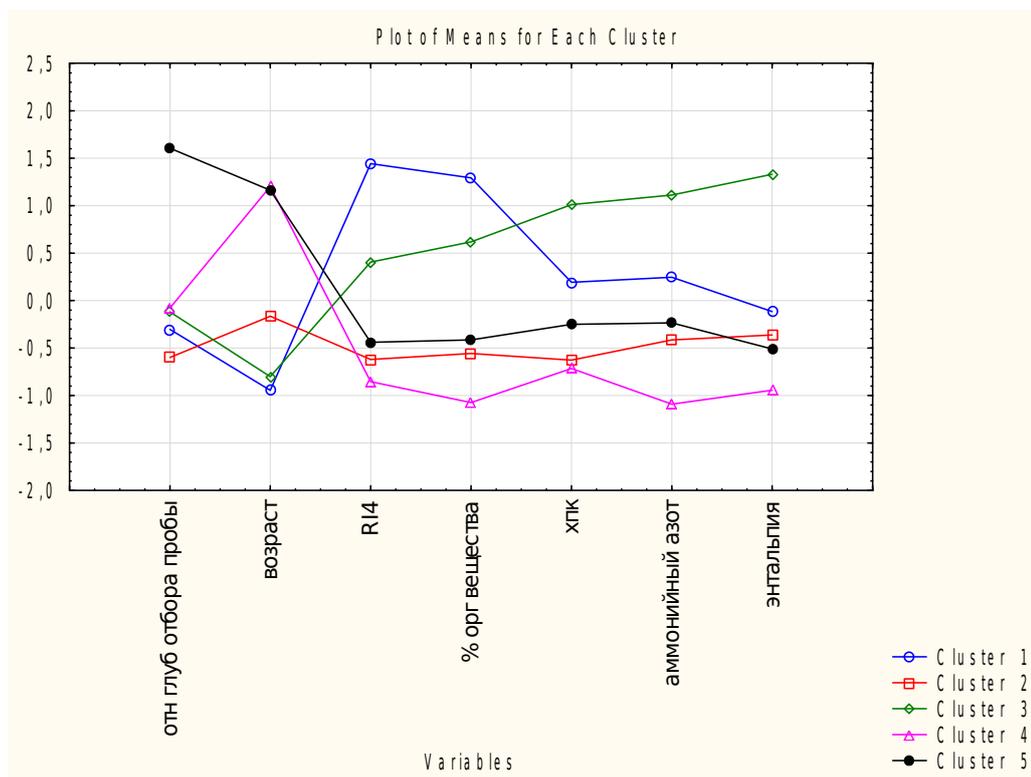


Рисунок 3.4 График средних значений показателей по 5 кластерам

Обсуждение результатов разбиения данных (см. Приложение Д):

- к кластеру № 1 относятся 9 проб отходов, взятые с самых молодых объектов с наибольшими значениями показателей теста RI4 и процента содержания органического вещества в пробе. К данному кластеру относится городская свалка ТБО г. Кунгур (эксплуатируемая);
- кластер № 2 содержит 10 наблюдений и характерен для объектов возраста 10-30 лет с невысокими показателями экологического мониторинга. Второму кластеру принадлежит половина проб, взятых со свалки ТБО г. Краснокамск;

- кластер № 3 содержит 14 наблюдений и характерен для молодых полигонов ТБО. Пробы отходов взяты с полигонов с высокими показателями ХПК, аммонийного азота и энтальпии. Сюда относится полигон ТБО «Верхняя Пышма»;
- кластер № 4 содержит информацию о 11 пробах отходов с самых старых полигонов, отходы на которых приняли стабильное состояние по данным экологического мониторинга. Это пробы отходов, взятые с неэксплуатируемой городской свалки ТБО г. Кунгур и свалки «Голый Мыс»;
- к кластеру № 5 относятся 7 проб отходов с самых старых полигонов, но показатели экологического мониторинга не приняли минимальные значения. К этому кластеру относятся части проб, взятые со свалки ТБО г. Краснокамск и полигона ТБО «Дубна Левобережная».

Определить какая кластеризация лучше можно по данным дисперсионного анализа. Чем меньше внутригрупповая дисперсия и больше межгрупповая, тем лучше рассматриваемый признак характеризует принадлежность объектов к кластеру и тем качественней кластеризация. В данном случае разбиение данных на 5 кластеров значительно лучше.

Исходя из данных кластеризации можно сделать вывод о том, что неэксплуатируемая часть городской свалки ТБО г. Кунгур и свалка «Голый Мыс» являются стабильными объектами захоронения твердых бытовых отходов. В разбиении данных на 3 и 4 кластера к группе, отвечающей за стабильные объекты, относится также свалка ТБО г.

Краснокамск, но некоторые пробы, взятые с этой свалки, были отнесены к другим кластерам. Следовательно, данную свалку нельзя считать стабильной.

То, что некоторые полигоны были отнесены к разным кластерам, обусловлено послойным складированием мусора. Получается, что в нижних слоях массива полигона содержатся старые отходы, и для них химико-биологические показатели уже пришли в стабильное состояние, когда как в верхних слоях содержатся молодые отходы и их показатели, соответственно, выше нормы.

Выявлено, что с возрастом все химико-биологические показатели экологического мониторинга уменьшаются и приходят в стабильное состояние с окружающей средой. Следовательно, анализируя данные экологического мониторинга с помощью кластерного анализа, можно получить разбиение объектов по возрасту и степени стабильности отходов на месте захоронения ТБО.

3.3. Исследование изменения фракционного состава

Отходы, которые располагаются на свалках и полигонах, подвергаются микробиологическому разложению и действию физико-механических факторов. Это приводит к тому, что морфологический и фракционный составы отходов изменяются [26].

Определим, зависит ли фракционный состав отходов от возраста полигона и от глубины отбора пробы. За примеры

объектов захоронения возьмем эксплуатируемую часть городской свалки ТБО г. Кунгур (представитель молодого объекта - 5 лет), полигон ТБО Верхняя Пышма (его возраст составляет 20-30 лет) и свалку ТБО «Голый Мыс» (закрытый полигон возраста 50 лет). Исследуемая таблица данных содержит информацию о 29 пробах отходов, взятых с трех полигонов, которые характеризуются возрастом отходов, глубиной отбора пробы и процентными содержаниями фракций различных размеров (< 20, 20-50, 50-100, > 100 мм) (см. Приложение 6).

Построим корреляционную матрицу, в которой отражены зависимости размеров фракций отходов от их возраста:

Таблица 3.7

Корреляционная матрица

Correlations (Spreadsheet35)							
Marked correlations are significant at p < .05000							
N=29 (Casewise deletion of missing data)							
Variable	Means	Std.Dev.	возраст	<20	20-50	50-100	>100
возраст	15,2244	19,4565	1,00000	0,89611	-0,51236	-0,70679	-0,74888
<20	48,9034	22,9302	0,89611	1,00000	-0,45180	-0,74187	-0,88960
20-50	12,0275	5,7615	-0,51236	-0,45180	1,00000	0,70211	0,02541
50-100	11,5282	5,9059	-0,70679	-0,74187	0,70211	1,00000	0,39393
>100	27,5382	17,9288	-0,74888	-0,88960	0,02541	0,39393	1,00000

По данным корреляционной матрицы можно сделать вывод о сильной зависимости между показателями возраста и процентом содержания мелких фракций в теле полигона.

С помощью кластерного анализа разобьем наши данные на 3 группы, которые характерны для различного фракционного состава отходов.

Таблица 3.8

Дисперсионный анализ

Variable	Analysis of Variance (fraktsost)					
	Between SS	df	Within SS	df	F	signif. p
возраст	33,4643	2	16,5356	48	48,5704	0,00000
<20 mm	34,7825	2	15,2175	48	54,8565	0,00000
20-50 mm	24,9647	2	25,0352	48	23,9324	0,00000
50-100 mm	28,0615	2	21,9384	48	30,6984	0,00000
>100 mm	29,7775	2	20,2224	48	35,3399	0,00000

По данным дисперсионного анализа видно, что разбиение качественно, о чем свидетельствуют большие значения критерия Фишера и маленькие значения p .

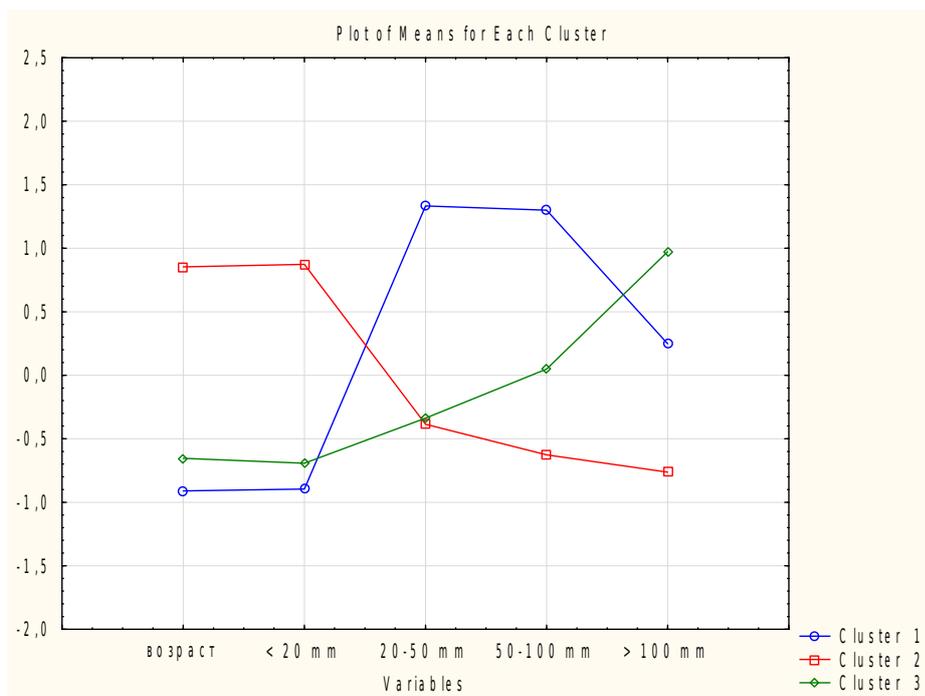


Рисунок 3.5 График средних значений параметров

Интерпретируя данные кластерного анализа можно сделать вывод о том, что с увеличением возраста объекта захоронения ТБО растет содержание в нем мелких фракций размером $< 20\text{mm}$ и уменьшается содержание крупных фракций.

Построим гистограммы, показывающие характер фракционного состава на полигоне.

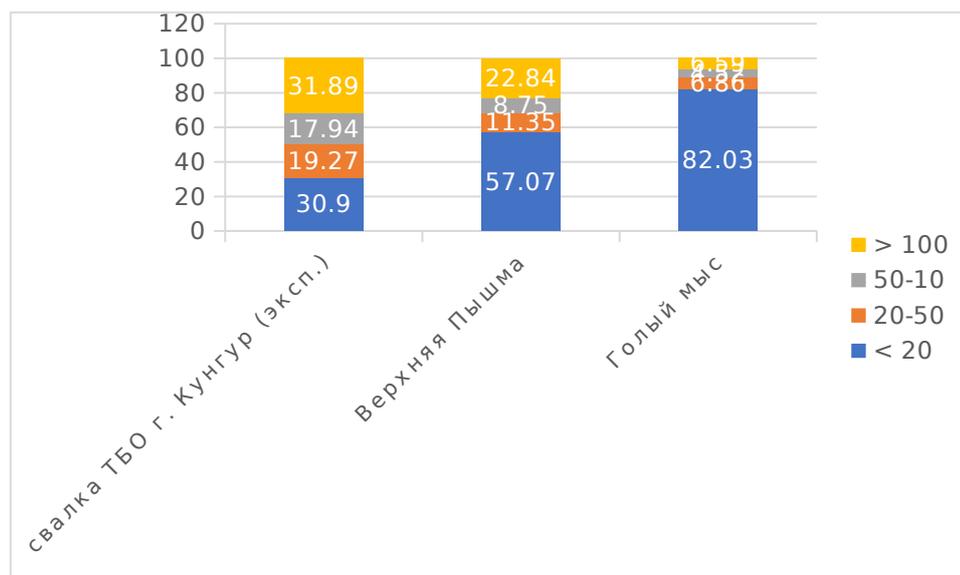


Рисунок 3.6 Фракционный состав ТБО

Сравнивая усредненные данные фракционного состава проб отходов, подтверждается вышесказанный вывод. Действительно, содержание мелких фракций на старой свалке «Голый Мыс» в 2,6 раз больше, чем их содержание на молодой свалке. Это связано с тем, что отходы более длительное время подвергались воздействиям биохимических и физико-механических факторов. Тело полигона создает благоприятные условия для микробиологического разложения биоразлагаемых отходов. Так же отходы на полигоне подвергаются деформации, слеживаемости, а раз увеличение времени воздействия различных биохимических процессов влияет на увеличение хрупкости объекта, ослабляет сопротивляемость физико-механическим изменениям, это ведет к уменьшению содержания крупных фракций отходов.

Построим гистограммы по трем объектам исследования, отражающие изменение фракционного состава от глубины отбора пробы и посмотрим, зависят ли они друг от друга.

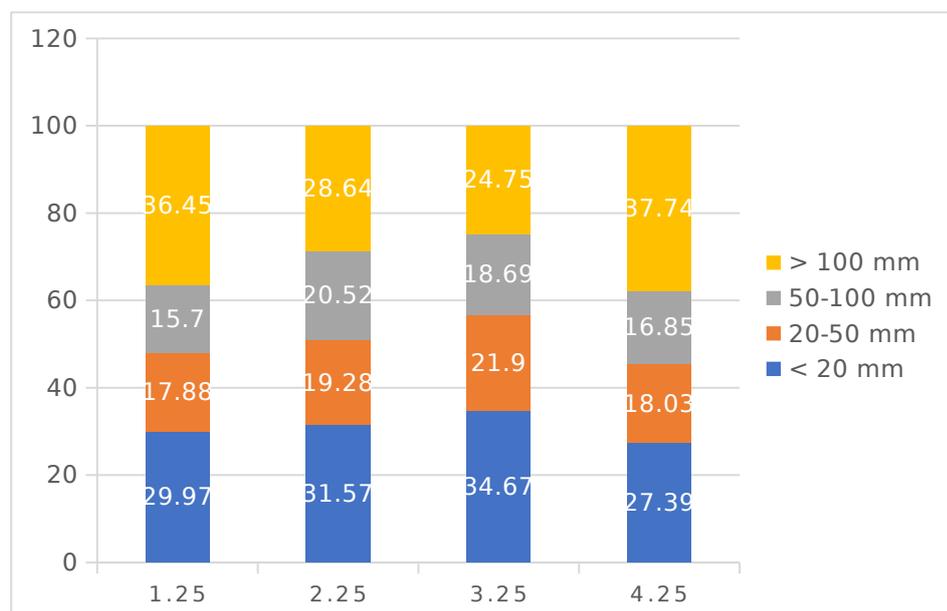


Рисунок 3.7 Фракционный состав эксплуатируемого участка свалки ТБО г. Кунгур по глубине

Отходы, которые захоронены на молодом полигоне, не имеют закономерностей изменения фракционного состава. Это объясняется тем, что ТБО еще незначительно подверглись различным биохимическим и физико-механическим процессам.

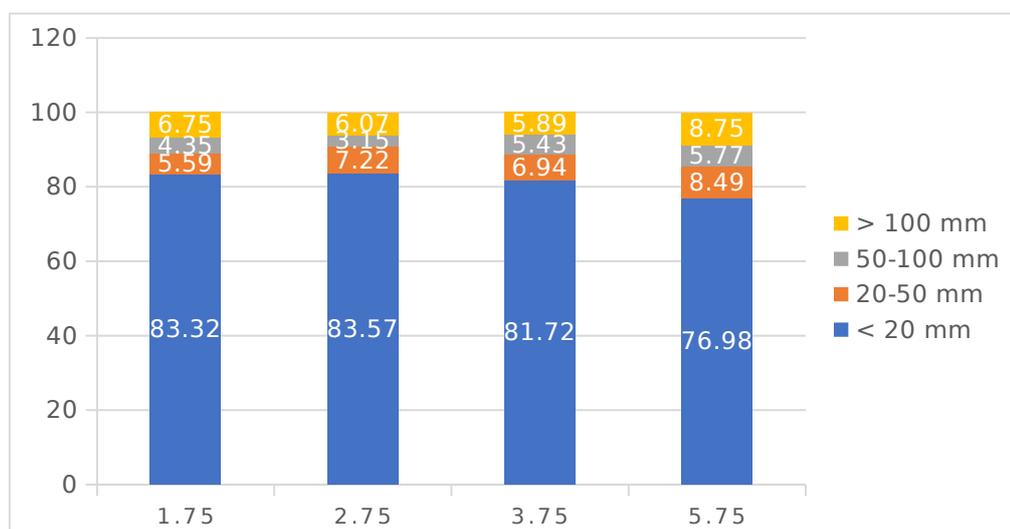


Рисунок 3.8 Фракционный состав свалки ТБО «Голый Мыс» по глубине

Глядя на визуализацию данных по свалке «Голый Мыс» можно сделать вывод о том, что в старых рекультивированных полигонах так же нет тенденции к изменению фракционного состава при изменении глубины отбора пробы. Одинаковое содержание фракций отходов по глубине зависит от того, что со временем выравнивается степень воздействия физико-механических и биохимических факторов.

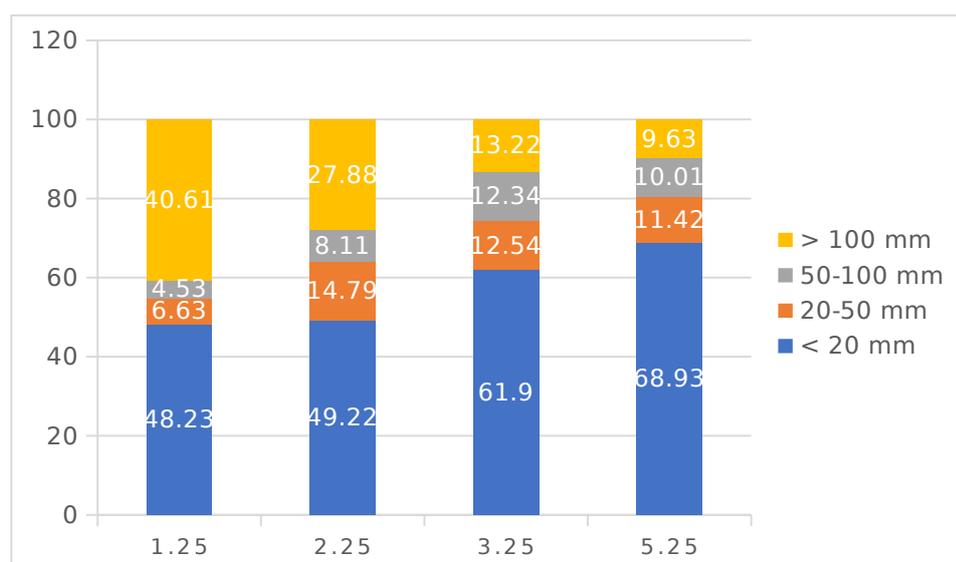


Рисунок 3.9 Фракционный состав полигона ТБО «Верхняя Пышма» по глубине

Если рассматривать полигон ТБО «Верхняя Пышма», то видно, что с увеличением глубины отбора пробы растет и содержание фракций размером <20 мм. Это объясняется тем, что отходы на места захоронения поступают неравномерно и складываются слоями. Следовательно, в нижних слоях

массива полигона содержатся более старые отходы, чем наверху, а как раньше выяснилось, содержание мелких фракций на свалке напрямую зависит от возраста захоронения. Воздействие физико-механических факторов на структуру ТБО увеличивается с течением времени и по глубине отбора пробы, следовательно, процессы деформации более крупных фракций протекают с большей интенсивностью в глубоких слоях массива полигона.

3.4. Исследование изменения морфологического состава

После определения фракционного состава по пробам отходов каждую фракцию отходов разбирают вручную с целью определения морфологического состава ТБО на исследуемом объекте. Были получены данные о содержании в пробе таких предметов как стекло, камни, металл, древесина, полимеры, текстиль, бумага, свалочный грунт (мелкая неидентифицируемая масса размером менее 2 мм) (см. Приложение Е).

Так же рассмотрим три объекта разного возраста: эксплуатируемый участок городской свалки ТБО г. Кунгур, полигон ТБО «Верхняя Пышма» и закрытую свалку ТБО «Голый Мыс». Разобьем данные морфологического состава на 4 кластера и проверим, как изменяется морфологический состав ТБО в зависимости от возраста.

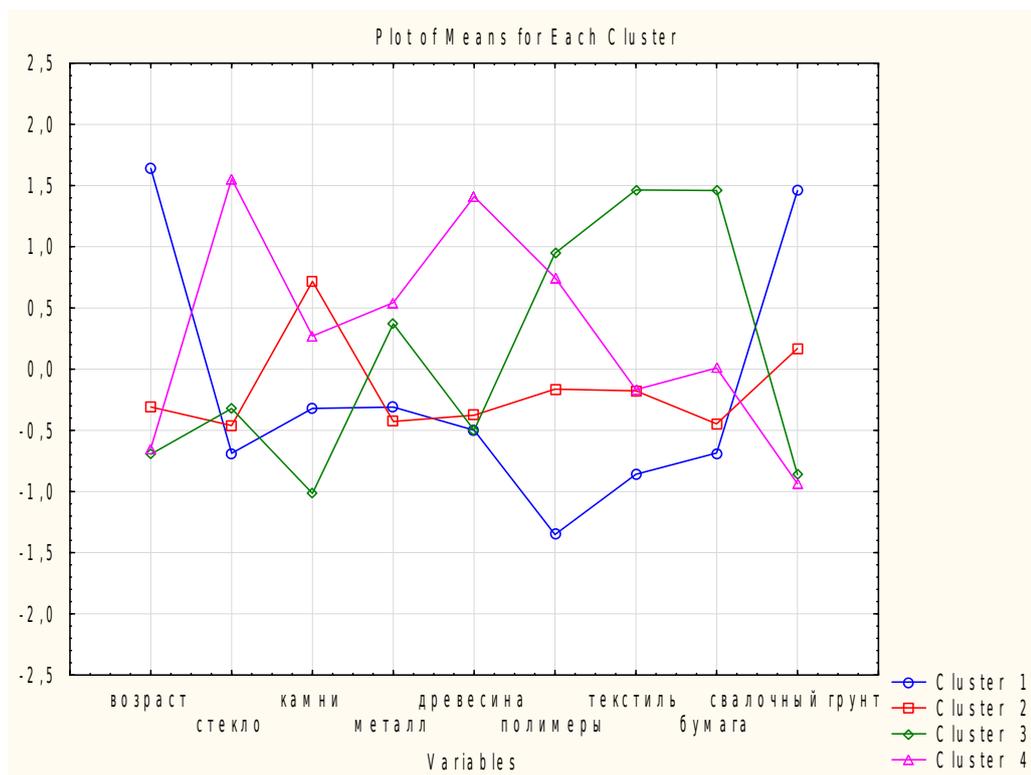


Рисунок 3.10 График средних значений параметров

Анализируя данные графика средних значений по каждому кластеру, можно сделать следующие выводы:

- чем старше отходы, тем меньше в них содержание биоразлагаемых материалов, таких как древесина, бумага и текстиль;
- для старых объектов характерно большое содержание свалочного грунта в теле полигона, что объясняется большей степенью разложения биodeградируемых материалов и их переходом в состав свалочного грунта.

Так же отмечено, что для старых закрытых свалок характерны низкие содержания неразлагаемых материалов (стекло, камни, металл, пластмасса). Это объясняется тем, что состав выбрасываемого мусора значительно отличается по времени. 20 лет назад не было одноразовой посуды, такого огромного количества полиэтиленовых пакетов, бумажной

обертки. Люди мало выбрасывали стеклянную посуду и бумагу.

- Кластеру № 1 принадлежит рекультивированная свалка ТБО «Голый Мыс». Она характерна высоким содержанием свалочного грунта в массиве полигона, что обусловлено возрастом отходов;
- полигон ТБО «Верхняя Пышма» относится к кластерам 2 и 3. Такой разброс по кластерам обусловлен тем, что нижние слои отходов соответствуют возрасту 20-30 лет (меньше содержание биоразлагаемых материалов и большое содержание свалочного грунта), а верхние 0-1 год (большое содержание биоразлагаемых материалов и низкие значения содержания свалочного грунта);
- эксплуатируемая часть городской свалки ТБО г. Кунгур принадлежит кластеру № 3. Для самой молодой свалки характерны высокие показатели содержания неразлагаемых материалов и низкие содержания свалочного грунта по всему массиву полигона.

3.5.Факторный анализ экологического мониторинга полигонов ТБО

Рассмотрим данные о химико-биологических показателях, морфологическом и фракционном составе. Построим регрессию, позволяющую по данным экологического мониторинга определить степень стабильности полигонов и свалок твердых бытовых отходов. Степень стабильности объекта захоронения определяется по

данным кластерного анализа по химико-биологическим показателям. То есть всего 5 степеней стабильности. Самые молодые свалки имеют коэффициент стабильности 1, самые старые и стабильные полигоны имеют коэффициент стабильности 5.

Для анализа используем следующие химико-биологические показатели: тест на дыхательную активность (RI4), тест на газообразование (GS21), процент содержания органического вещества в пробе, ХПК, БПК, аммонийный азот, энтальпия.

Из данных о морфологическом составе возьмем показатели содержания полимеров, текстиля, бумаги и свалочного грунта. Показатели содержания неразлагаемых материалов в исследовании не участвуют, так как их содержание практически не зависит от возраста отходов.

Из данных об фракционном составе возьмем все показатели.

Перед тем, как проводить факторный анализ проведем тесты КМО и Бартлетта на определение адекватности применения факторизации к данным.

Таблица 3.9

КМО и критерий Бартлетта

Мера адекватности выборки Кайзера-Майера-Олкина (КМО).		0,743
Критерий сферичности Бартлетта	Примерная Хи-квадрат	1172,605
	ст.св.	105
	Значимость	0,000

Значение критерия Кайзера-Майера-Олкина получилось равным 0,74, это значит, что адекватность применения факторного анализа находится на приемлемом уровне. Следовательно, факторный анализ использовать корректно.

Из показателей критерия сферичности Бартлетта также следует адекватность применения факторного анализа.

Используем самый распространенный метод факторного анализа – метод главных компонент.

Метод главных компонент в программе STATISTICA выделил среди всех показателей три фактора (главные компоненты).

Рассмотрим матрицу собственных значений:

Таблица 3.10

Матрица собственных значений

Значение	Собственные значения	% общей дисперсии	Кумулятивные собственные значения	Кумулятивный %
1	8,99959	56,24	8,99959	56,24749
2	2,34718	14,66	11,3467	70,91738
3	1,07581	6,723	12,4226	77,64123

На основании собственных чисел можно определить вклад главных компонент в суммарную дисперсию системы. Суммарная дисперсия системы равна 15 (количество исходных переменных). Получилось, что первая компонента объясняет 56% от общей дисперсии, а все три фактора вместе объясняют 77,6% суммарной дисперсии.

На основе матрицы факторных нагрузок дадим объяснение полученных главных компонент.

Таблица 3.11

Матрица факторных нагрузок

Переменная	Фактор 1	Фактор 2	Фактор 3
RI4	<u>0,775</u>	0,253	0,591
GS21	<u>0,718</u>	0,276	0,423
% органики	<u>0,86</u>	0,443	0,534
ХПК	0,045	0,475	<u>0,792</u>
БПК	0,397	0,033	<u>0,787</u>
аммонийный азот	0,17	0,491	<u>0,703</u>
Энтальпия	0,092	<u>0,731</u>	0,324
полимеры	0,57	0,567	0,322
текстиль	-0,104	0,647	0,552
бумага	0,204	<u>0,791</u>	0,124
свалочный грунт	-0,709	-0,623	-0,262
<20	0,633	<u>-0,71</u>	-0,201
20-50	0,622	-0,157	-0,122
50-100	0,68	0,292	0,138
>100	0,251	<u>0,851</u>	0,249

Получилось, что фактор 1 сильно коррелирует (зависит) с такими переменными, как тесты RI4 и GS21 и процент содержания органического вещества в пробе. Таким образом, этот фактор можно интерпретировать, как способность к биологическому разложению.

Фактор 2 зависит от энтальпии, содержания бумаги в пробе и процента содержания крупных фракций. Энтальпия показывает, сколько энергии доступно для преобразования в теплоту, следовательно, этот фактор можно интерпретировать как энергетический потенциал.

Фактор 3 коррелирует с показателями ХПК, БПК и аммонийный азот. Значит этот фактор можно обозначить как показатель загрязненности сточных вод.

Построим регрессию, которая определяет степень стабильности отходов в зависимости от наших новых факторов.

Таблица 3.12

Регрессионный анализ

Regression Summary for Dependent Variable: стабильность (Spreadsheet)						
R = .88078142 R ² = .77577590 Adjusted R ² = .76146372						
F(3,47) = 54.2044 p < 0.00000 Std. Error of estimate = .69354						
	b*	Std. Err. of b*	β	Std. Err. of β	t(47)	p-value
Intercept			2.941174	0.097113	30.2853	0.00000
F1	-.068385	0.069070	-.097108	0.098082	-9.9008	0.00000
F2	-.39159	0.069070	-.55607	0.098082	-5.6694	0.00000
F3	-.39341	0.069070	-.55866	0.098082	-5.6958	0.00000

Мы получили регрессию:

$$y=2,941-0.971*F_1-0,556*F_2-0.559*F_3, \quad (3.1)$$

где y – степень стабильности отходов;

F_1 – способность к биологическому разложению;

F_2 – энергетический потенциал;

F_3 – показатель загрязненности сточных вод.

Коэффициент детерминации полученной модели равен 0,776, следовательно, наша регрессия объясняет 77% исходных данных. Модель получилась значима.

Применяя данную регрессию к нашим данным, получилось, что максимальный коэффициент стабильности отходов (больше 4,5) принадлежит таким объектам, как закрытый участок городской свалки ТБО г. Кунгур и закрытая свалка ТБО «Голый Мыс».

После проведения исследования выяснилось, что эти два объекта относятся к стабильным полигонам. Однако, раз закрытый участок свалки г. Кунгур находится в непосредственной близости с эксплуатируемым участком, то его можно после рекультивации только использовать повторно в качестве нового места складирования ТБО. Территория свалки ТБО «Голый Мыс» также может использоваться повторно как новый полигон захоронения ТБО или же для другого предназначения.

ЗАКЛЮЧЕНИЕ

При подготовке дипломной работы была определена цель - найти зависимость степени стабильности отходов на объектах захоронения ТБО от данных об экологическом мониторинге с помощью методов многомерного статистического анализа.

Реализации цели дипломного исследования способствовало рассмотрение ряда задач, которые были решены следующим образом.

Твердые бытовые отходы представляют собой предметы, которые человек уже использовал, надобность в которых утрачена. Это тот мусор, который люди ежедневно выкидывают. Все ТБО поступают на специализированные площадки - полигоны твердых бытовых отходов. Мусор, который остался после сортировки ТБО, отправляется на захоронение в массив полигона. Каждый полигон имеет свой срок эксплуатации и по его истечению закрывается и рекультивируется с целью дальнейшего использования территории.

На каждом полигоне ТБО обязан проводится экологический мониторинг, в ходе которого определяется степень стабильности отходов по данным химико-биологических показателей, изучается их воздействие на окружающую среду, определяется фракционный и морфологический состав ТБО.

Методы многомерного статистического анализа позволяют работать с большими совокупностями данных. Кластерный анализ позволяет разбить данные на группы схожих между собой объектов, факторный анализ позволяет сократить размерность исследуемых показателей, а регрессионный анализ позволяет представить изменение результирующей переменной в виде линейной регрессии.

На основании проделанного исследования можно сделать следующие выводы:

- степень стабильности полигонов и свалок ТБО зависит от возраста объекта и значений химико-биологических показателей. Чем старше захороненные отходы, тем меньше значения химико-биологических показателей, следовательно, старые закрытые полигоны с наименьшими показателями экологического мониторинга являются стабильными;
- фракционный состав отходов меняется в зависимости от их возраста. Чем старше полигон, тем больше в нем содержание мелких фракций размером меньше 20 мм и меньше содержание крупных фракций размером больше 100 мм. В полигонах, возраст отходов на которых составляет 0-30 лет, отмечена зависимость содержания мелких фракций от глубины отбора пробы: чем глубже отходы, тем больше в них содержание мелких фракций;
- морфологический состав ТБО так же изменяется с течением времени. Чем больше возраст отходов, тем меньше в нем содержание биоразлагаемых материалов и больше содержание свалочного грунта.

Так же в ходе дипломной работы была построена регрессия, позволяющая определить степень стабильности отходов на полигонах по данным экологического мониторинга, который, в свою очередь, можно представить в виде трех главных факторов.

Данная дипломная работа отличается высокой практической значимостью. В ходе ее создания было установлено, что с помощью методов многомерного статистического анализа можно изучать данные о полигонах и свалках ТБО.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Harman G. Modern factor analysis. – М.: Statistika, 1972. – 484 с.
2. Huber-Humer M., Roeder S., Lechner P. Approaches to assess biocover performance on landfills // Waste Management 29. – 2009. – P. 2092–2104.
3. Rencher A. C. Methods of multivariate analysis. Wiley. 2002. 739 p.
4. Soyez K., Plickert S. (2002) Mechanical-Biological Pretreatment of Waste: State of the art and potentials of biotechnology. Acta Biotechnologica. – Vol. 22, № 3–4. – P. 271–284.
5. Буреева Н.Н. Многомерный статистический анализ с использованием ППП «STATISTICA» / Н.Н. Буреева. — Н. Новгород: ННГУ, 2007. — 112с.
6. Гитис Л.Х. Факторный анализ: Лекция 1. Постановка задачи и области применения / Л.Х. Гитис. – М.: Издательство Московского государственного горного университета, 2000. – 24 с.
7. Завизион Ю.В. Оценка состояния полигонов захоронения твердых коммунальных отходов на этапах жизненного цикла – термоаналитический и спектроскопический подход / Пермский национальный исследовательский политехнический университет. – Пермь: ПНИПУ, 2016. – 20 с.
8. Изучение состава отходов массива полигона ТБО / Ю.М. Загорская [и др.] // Вестник ПНИПУ. Урбанистика. – 2013. – №4. – с. 144-154

9. Калинина В.Н., Соловьев В.И. Введение в многомерный статистический анализ: Учебное пособие / В.Н. Калинина, В.И. Соловьев. – М.: ГУУ, 2003. – 66с.
10. Кит П. Управленческая экономика: инструментарий руководителя = Managerial Economics: economic tools for today's decision makers / П. Кит, Ф. Янг ; пер. с англ. : Л. Круглов-Морозов, И. Малкова. - 5-е изд. - Санкт-Петербург : Питер, 2009. - 622 с.
11. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников / А.И. Кобзарь. – М.: ФИЗМАТЛИТ, 2006.
12. Линник Ю.В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений / Ю.В. Линник. – М.: Физматгиз, 1958. – 336 с.
13. Лоули Д. Факторный анализ как статистический метод / Д. Лоули, А. Максвелл. – М.: Мир, 1967. – 137 с.
14. Мандель И. Д. Кластерный анализ / И.Д. Мандель. — М.: Финансы и статистика. 1988. — 176 с.
15. Машин В.А. Методические вопросы использования факторного анализа на примере спектральных показателей сердечного ритма // Экспериментальная психология. – 2010. – т. 3, №4. – с. 119-138
16. Многомерный статистический анализ в экономике: Учеб. пособие для вузов / Под ред. проф. В. Н. Тамашевича. — М.: ЮНИТИ-ДАНА, 1999. — 598 с.
17. Мокеев В.В. Об использовании метода главных компонент для анализа деятельности предприятия / В.В. Мокеев, К.Л. Соломахо. // Вестник ЮУрГУ. Серия «Экономика и менеджмент». – 2013. – т. 7, № 3. – с. 41-46

18. Очистка сточных вод предприятий при высоком ХПК и БПК [Электронный ресурс] // URL: <http://www.osmotics.ru/ru/vse-o-stokah/podrobno-pro-hpk-i-bpk>
19. Практикум по эконометрике. А.В. Гладилин, А.Н. Герасимов, Е.И. Громов // Ростов н/Д: Феникс, 2011. – 326 с.
20. Сайт о переработке и утилизации отходов [Электронный ресурс] // URL: <https://vtorothodi.ru/>
21. Состояние и перспективы российского рынка переработки твердых бытовых отходов. М.: Техноконсалт, 2003. – 73 с.
22. Статистика: учебник для академического бакалавриата / отв. ред. И.И. Елисеева. – 5-е изд., пер. и доп. – М. : Издательство Юрайт, 2019. – 572 с. – (Серия: Бакалавр. Академический курс).
23. Челноков А.А. Обращение с отходами. Для студентов учреждений высшего образования / А.А. Челноков, Л.Ф. Ющенко, И.Н. Жмыхов, К.К. Юращик. – Минск: «Вышэйшая школа», 2018. – 447 с.
24. Что такое ТБО и что обязательно нужно знать [Электронный ресурс] // URL: <https://vtorothody.ru/othody/chto-takoe-tbo.html>
25. Шубов Л.Я. Технологии отходов (Технологические процессы в сервисе): Учебник / Л.Я. Шубов, М.Е. Ставровский, Д.В. Шехирев. – М.: ГОУВПО «МГУС», 2006. – 410 с.
26. Энтальпия [Электронный ресурс] // URL: <http://www.xiron.ru/content/view/23194/28/>

ПРИЛОЖЕНИЕ А

Химико-биологические показатели

Таблица А.1

Пример данных по химико-биологическим показателям

Скважина	Свалка "Голый мыс"						
	Скважина №1			Скважина №2			
Высота массива полигона, м	8	8	8	8	8	8	8
Глубина	1,5-2 м	2,5-3 м	4,5-5 м	1,5-2 м	2,5-3 м	4,5-5 м	5,5-6 м
Относительная глубина отбора пробы, м	0,2	0,3	0,6	0,2	0,3	0,6	0,7
Средний возраст, лет	40	50	50	40	50	50	50
Дыхательная активность, мгО ₂ /г (сух. в-ва)	0,6	0,5	1,0	0,3	0,7	2,3	0,3
Органическое вещ-во, %	9	10	9	16	17	26	22
ХПК, мг О ₂ /л	265	150	147	237	140	135	115
БПК ₅ , мг О ₂ /л	20	5	5	10	6,7	6,7	5
Зольность, %	91	90	91	84	83	74	78
N-NH ₄ , мг/л	65	58	58	40	58	177	449
GS21, л/кг сух. в-ва	1,9	1,5	4,9	0,6	1,8	5,6	1,9
Энтальпия, Дж/г	905	669	396,9	1007	1678	681,8	1723

ПРИЛОЖЕНИЕ Б

Корреляционная матрица исходных данных

Таблица Б.1

Корреляционная матрица исходных данных

Correlations (TBOish) Marked correlations are significant at $p < ,05000$ N=51 (Casewise deletion of missing data)										
Variable	отн глуб отбора пробы	возраст	R4	% орг вещества	хпк	бпк	зольность	аммонийный азот	GS21	энтальпия
отн глуб отбора пробы	1,00000	0,43707	-0,16563	-0,20147	-0,13692	-0,18582	0,20147	-0,00043	-0,22709	-0,13254
возраст	0,43707	1,00000	-0,63484	-0,74536	-0,48424	-0,51733	0,74536	-0,59598	-0,61214	-0,70201
R4	-0,16563	-0,63484	1,00000	0,83516	0,61518	0,71173	-0,83516	0,51609	0,92517	0,39607
% орг вещества	-0,20147	-0,74536	0,83516	1,00000	0,71810	0,60579	-1,00000	0,62458	0,78997	0,56453
хпк	-0,13692	-0,48424	0,61518	0,71810	1,00000	0,56891	-0,71810	0,71025	0,46600	0,58307
бпк	-0,18582	-0,51733	0,71173	0,60579	0,56891	1,00000	-0,60579	0,43315	0,59568	0,36865
зольность	0,20147	0,74536	-0,83516	-1,00000	-0,71810	-0,60579	1,00000	-0,62458	-0,78997	-0,56453
аммонийный азот	-0,00043	-0,59598	0,51609	0,62458	0,71025	0,43315	-0,62458	1,00000	0,38494	0,61331
GS21	-0,22709	-0,61214	0,92517	0,78997	0,46600	0,59568	-0,78997	0,38494	1,00000	0,38528
энтальпия	-0,13254	-0,70201	0,39607	0,56453	0,58307	0,36865	-0,56453	0,61331	0,38528	1,00000

ПРИЛОЖЕНИЕ В

Состав кластеров при К= 3

Таблица В.1.

Кластер № 1

4	городская свалка ТБО Кунгур (эксп)
8	городская свалка ТБО Кунгур (эксп)
13	свалка ТБО Краснокамск
17	свалка ТБО Краснокамск
21	свалка ТБО Краснокамск
38	Верхняя Пышма ТБО
39	Верхняя Пышма ТБО
40	Верхняя Пышма ТБО
41	Верхняя Пышма ТБО
42	Верхняя Пышма ТБО
44	Верхняя Пышма ТБО
45	Верхняя Пышма ТБО
48	Дубна Левобережная
49	Дубна Левобережная
51	Дубна Левобережная

Таблица В.2.

Кластер № 2

9	городская свалка ТБО Кунгур
10	городская свалка ТБО Кунгур
11	городская свалка ТБО Кунгур
12	городская свалка ТБО Кунгур
14	свалка ТБО Краснокамск
15	свалка ТБО Краснокамск
18	свалка ТБО Краснокамск
19	свалка ТБО Краснокамск
20	свалка ТБО Краснокамск
22	свалка ТБО Краснокамск

23	свалка ТБО Краснокамск
24	свалка ТБО Краснокамск
25	свалка "Голый Мыс"
26	свалка "Голый Мыс"
27	свалка "Голый Мыс"
28	свалка "Голый Мыс"
29	свалка "Голый Мыс"
30	свалка "Голый Мыс"
31	свалка "Голый Мыс"
50	Дубна Левобережная

Таблица В.3.

Кластер № 3

1	городская свалка ТБО Кунгур (эксп)
2	городская свалка ТБО Кунгур (эксп)
3	городская свалка ТБО Кунгур (эксп)
5	городская свалка ТБО Кунгур (эксп)
6	городская свалка ТБО Кунгур (эксп)
7	городская свалка ТБО Кунгур (эксп)
32	Верхняя Пышма ТБО
33	Верхняя Пышма ТБО
34	Верхняя Пышма ТБО
35	Верхняя Пышма ТБО
36	Верхняя Пышма ТБО
37	Верхняя Пышма ТБО
43	Верхняя Пышма ТБО
46	Дубна Левобережная
47	Дубна Левобережная

ПРИЛОЖЕНИЕ Г

Состав кластеров при K= 4

Таблица Г.1.

Кластер № 1

4	городская свалка ТБО Кунгур (эксп)
8	городская свалка ТБО Кунгур (эксп)
13	свалка ТБО Краснокамск
14	свалка ТБО Краснокамск
17	свалка ТБО Краснокамск
18	свалка ТБО Краснокамск
21	свалка ТБО Краснокамск
22	свалка ТБО Краснокамск
23	свалка ТБО Краснокамск
41	Верхняя Пышма ТБО
42	Верхняя Пышма ТБО
44	Верхняя Пышма ТБО
45	Верхняя Пышма ТБО
48	Дубна Левобережная
49	Дубна Левобережная
51	Дубна Левобережная

Таблица Г.2.

Кластер № 2

1	городская свалка ТБО Кунгур (эксп)
2	городская свалка ТБО Кунгур (эксп)
3	городская свалка ТБО Кунгур (эксп)
5	городская свалка ТБО Кунгур (эксп)
6	городская свалка ТБО Кунгур (эксп)
7	городская свалка ТБО Кунгур (эксп)

Таблица Г.3.

Кластер № 3

32	Верхняя Пышма ТБО
33	Верхняя Пышма ТБО
34	Верхняя Пышма ТБО
35	Верхняя Пышма ТБО
36	Верхняя Пышма ТБО
37	Верхняя Пышма ТБО
38	Верхняя Пышма ТБО
39	Верхняя Пышма ТБО
40	Верхняя Пышма ТБО
46	Дубна Левобережная
47	Дубна Левобережная

Таблица Г.4.

Кластер № 4

9	городская свалка ТБО Кунгур
10	городская свалка ТБО Кунгур
11	городская свалка ТБО Кунгур
12	городская свалка ТБО Кунгур
15	свалка ТБО Краснокамск
16	свалка ТБО Краснокамск
19	свалка ТБО Краснокамск
20	свалка ТБО Краснокамск
24	свалка ТБО Краснокамск
25	свалка "Голый Мыс"
26	свалка "Голый Мыс"
27	свалка "Голый Мыс"
28	свалка "Голый Мыс"
29	свалка "Голый Мыс"
30	свалка "Голый Мыс"
31	свалка "Голый Мыс"
50	Дубна Левобережная

ПРИЛОЖЕНИЕ Д

Состав кластеров при K= 5

Таблица 5.1.

Кластер № 1

1	городская свалка ТБО Кунгур (эксп)
2	городская свалка ТБО Кунгур (эксп)
3	городская свалка ТБО Кунгур (эксп)
4	городская свалка ТБО Кунгур (эксп)
5	городская свалка ТБО Кунгур (эксп)
6	городская свалка ТБО Кунгур (эксп)
7	городская свалка ТБО Кунгур (эксп)
8	городская свалка ТБО Кунгур (эксп)
43	Верхняя Пышма ТБО

Таблица 5.2.

Кластер № 2

13	свалка ТБО Краснокамск
14	свалка ТБО Краснокамск
17	свалка ТБО Краснокамск
18	свалка ТБО Краснокамск
19	свалка ТБО Краснокамск
21	свалка ТБО Краснокамск
22	свалка ТБО Краснокамск
23	свалка ТБО Краснокамск
42	Верхняя Пышма ТБО
44	Верхняя Пышма ТБО

Таблица 5.3.

Кластер № 3

32	Верхняя Пышма ТБО
33	Верхняя Пышма ТБО
34	Верхняя Пышма ТБО
35	Верхняя Пышма ТБО
36	Верхняя Пышма ТБО
37	Верхняя Пышма ТБО
38	Верхняя Пышма ТБО
39	Верхняя Пышма ТБО
40	Верхняя Пышма ТБО
41	Верхняя Пышма ТБО
46	Дубна Левобережная
47	Дубна Левобережная
48	Дубна Левобережная
49	Дубна Левобережная

Таблица 5.4.

Кластер № 4

9	городская свалка ТБО Кунгур
10	городская свалка ТБО Кунгур
11	городская свалка ТБО Кунгур
12	городская свалка ТБО Кунгур
25	свалка "Голый Мыс"
26	свалка "Голый Мыс"
27	свалка "Голый Мыс"
28	свалка "Голый Мыс"
29	свалка "Голый Мыс"
30	свалка "Голый Мыс"
31	свалка "Голый Мыс"

Таблица 5.5.

Кластер № 5

15	свалка ТБО Краснокамск
16	свалка ТБО Краснокамск
20	свалка ТБО Краснокамск
24	свалка ТБО Краснокамск
45	Верхняя Пышма ТБО
50	Дубна Левобережная
51	Дубна Левобережная

ПРИЛОЖЕНИЕ Е

Морфологический и фракционный состав

Таблица Е.1

Пример данных о морфологическом и фракционном составе

	Свалка "Голый мыс"						
Скважина	Скважина №1			Скважина №2			
Высота массива полигона, м	8	8	8	8	8	8	8
Глубина	1,5-2 м	2,5-3 м	4,5-5 м	1,5-2 м	2,5-3 м	4,5-5 м	5,5-6 м
Относительная глубина отбора пробы, м	0,2	0,3	0,6	0,2	0,3	0,6	0,7
Средний возраст, лет	40	50	50	40	50	50	50
Стекло	2,1	2,3	2,9	4,7	4,4	6,2	4,0
Камни	9,6	12,5	8,6	12,2	9,4	6,2	12,0
Металл	0,4	0,3	0,3	2,1	2,0	2,4	1,2
Древесина	2,0	4,8	3,1	3,0	3,1	6,0	7,7
Полимеры	1,9	1,4	0,9	2,1	2,0	0,9	1,7
Текстиль	0,1	0,3	0,3	0,1	0,8	0,5	0,3
Бумага	0,1	0,1	0,1	0,1	0,1	0,7	0,5
Свалочный грунт	82,2	77,5	82,5	72,3	75,0	72,1	71,2
Прочее	1,7	0,9	1,5	3,4	3,4	5,0	1,3
< 20 мм	87,4	83,3	87,6	79,2	83,8	75,8	77,0
20-50 мм	3,5	5,6	4,7	7,7	8,8	9,2	8,5
50-100 мм	3,1	3,2	3,5	5,6	3,1	7,4	5,8
> 100 мм	6,0	7,9	4,2	7,5	4,2	7,6	8,8