



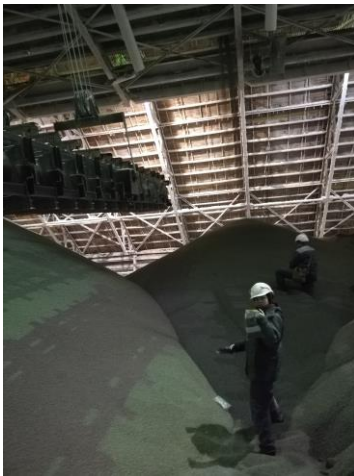
**МЕТОДЫ МНОГОМЕРНОЙ
КЛАССИФИКАЦИИ
ПРОМЫШЛЕННО
ПРОИЗВОДИМЫХ
МИНЕРАЛЬНЫХ УДОБРЕНИЙ
ДЛЯ ПРЕДСТАВЛЕНИЯ
МОДЕЛИ "ПРОИЗВОДИТЕЛЬ-
КАЧЕСТВО"**



Руководитель группы экспресс-методов анализа и автоматизации технологических процессов АО «НИУИФ», к.т.н. Юновидов Д.В.

Введение. Минеральные удобрения

Для роста и развития растений одинаково важны все элементы питания: при избытке одного и недостатке другого элемента получение высоких урожаев НЕВОЗМОЖНО.



Максимальное усвоение при соотношении элементов питания

N/S 10/2

P/Zn 10/0,3

K/Mg 10/1,5

Ca/B 10/0,1

Зерновые		
N	P	S
11	4	2

Кукуруза		
N	P	S
14	5	1

Картофель		
N	P	S
17	5	2

Сахарная свекла		
N	P	S
9	3	3

Минеральные удобрения – устойчивый и развивающийся рынок для развития сельского хозяйства и экосистемы (потребление $P_2O_5 > 39$ млн.т.). Необходимо обеспечить качество данного рынка.

Введение. Цель и задачи

Цель: построение модели «производитель-качество» для идентификации удобрений по брендовой и территориальной принадлежности.

Методы: энергодисперсионный рентгенофлуоресцентный анализ (ЭД РФА), понижение размерности данных и кластеризация методом k-means.

Задачи:

- Описание объекта исследования.
- Выбор способа и метода выражения качества.
- Обзор методов анализа «больших данных».
- Применение многомерной кластеризации для построения модели «производитель-качество».

Минеральные фосфорсодержащие удобрения

Свойства минеральных удобрений в первую очередь зависят от исходного сырья.

Таблица 1. Основные страны, добывающие фосфориты. Для удобства метрические тонны пересчитаны на тонны (данные ориентировочные).

Страна	Уровень добычи, 10^4 т. в год	Ориентировочные запасы, 10^7 т.
Китай	8,9	1
США	2,92	0,4
Марокко	2,8	2,1
Россия	1,13	0,1

- Каждое из месторождений обладает уникальным химическим составом, который зависит от типа фосфоритов (осадочные или магматические) и географического положения.
- Каждый из производителей зачастую использует уникальный состав сырья (или смесь) и процесс производства.

=> Уникальные свойства и химический состав произведенного продукта.

Качество минеральных удобрений

Характеристики, определяющие **качество продукции**: содержание питательных элементов (тип и марка, например NPK 15-15-15), растворимость, рассыпчатость, содержание микроэлементов (Zn, Mg и т.д.), содержание тяжелых элементов и т.д.

Для эффективной классификации продукции – **ЭД РФА**: широкий диапазон (0,0001 - 100 мас. %); простая пробоподготовка; экспрессность; многоэлементность; простота автоматизации и т.д.

1. Пробоподготовка

- 1 - 10 минут
- ~ 80 % погрешности

2. Измерение

- 5 - 300 секунд
- ~ 5 % погрешности

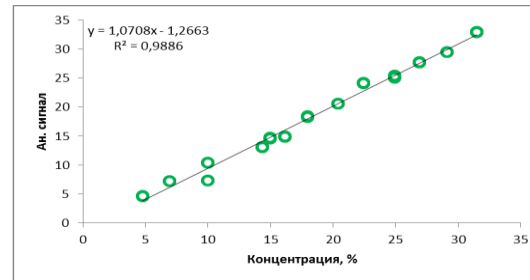
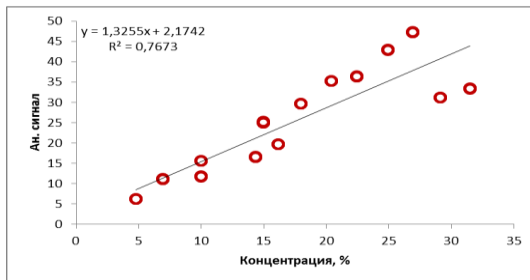
3. Расчет концентрации

- 1 - 60 секунд
- остальные 15 % погрешности



Методы многомерной классификации

Несмотря на развитость математического аппарата и разнообразие приборного парка, рассматриваемый метод не нашел широкого применения для сложных промышленных объектов.



Анализ больших данных (АБД) – это комплексная, многомерная и непрерывная статистическая обработка постоянно увеличивающихся объемов характеристической информации, для создания автоматизированных программ, алгоритмов и баз данных для объективного, своевременного и комплексного контроля.

Общие принципы АБД

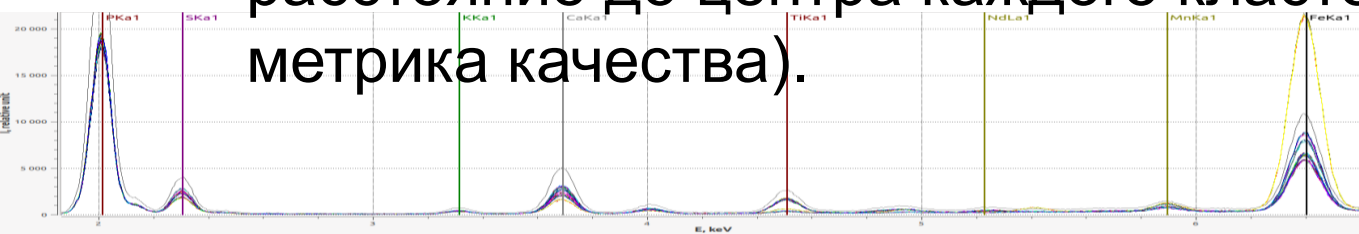
1. Создание матрицы данных «объекты-признаки».
2. Обработка данных в матрице (нормирование, удаление выбросов и линейно зависимых переменных) .
3. Использование статистических методов работы с данными: регрессия; классификация; понижение размерности данных и кластеризация.

Нет необходимости определять конкретные концентрации, достаточно использовать весь полученный спектр и провести кластеризацию полученной информации.

Модель «производитель-качество»

Основная идея - использование каждого канала спектра как признака (около 4000 каналов на один спектр).

1. Составляется матрица «объекты-признаки».
2. Признаки нормализуются по среднему и дисперсии.
3. Проводится понижение размерности до 2 компонент (плоскость) по алгоритму PCA.
4. Рассчитываются кластеры по алгоритму k-means и каждому объекту присваивается соответствующий номер.
5. Для каждого спектра рассчитывается критерий χ^2 и расстояние до центра каждого кластера (Евклидова метрика качества).



Выбранные метрики позволяют получить **количественную оценку величины класса и степени близости различных групп объектов** - внутреннее **качество продукции** по характерным наборам свойств.

Зона кластера, в которую попал неизвестный спектр с поправкой на его критерий χ^2 указывают на производителя. Расстояние до центра кластера является количественным критерием «качества» исследуемого удобрения.

Эксперимент

В работе исследован набор проб:

- Met* NPK(S) 4-30-15(16);
- Met* NPK(S) 0-20-20(5);
- Ара** NP(S) 12-40(10);
- Ара** NPK 16-16-8; Ара NP 12-52.

* Волховский филиал АО «Апатит», г. Волхов

** АО «Апатит», г. Череповец



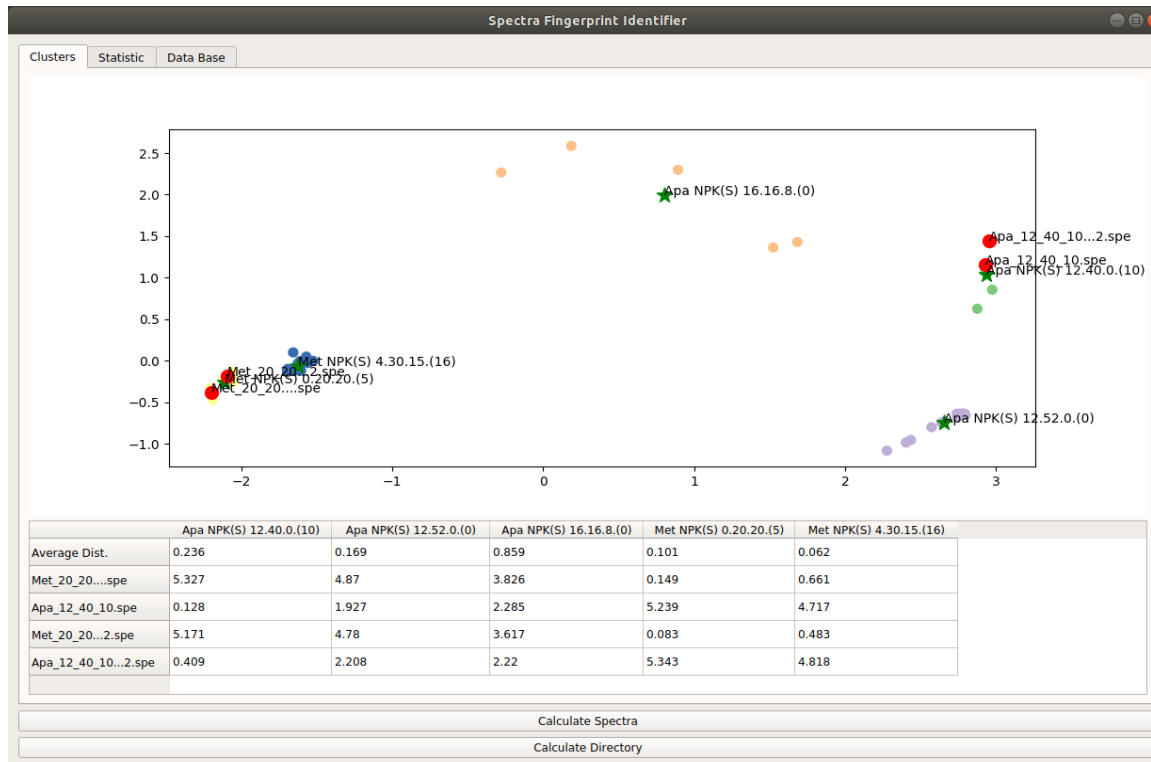
- Аналитическую пробу гранул (200 г) сокращали до 50 г, случайным образом отбирали по три навески гранул для формирования таблетки (2 г).
- Прессовали в виде «сэндвич-структуры» при 280 бар (11 т/см²) в течении 30 с.

В качестве маркеров используется весь спектр ЭД РФА анализа (учет явных и не явных свойств).

По результатам ЭД РФА проводится многомерный анализ образцов с целью их идентификации.

Результаты и обсуждение

Для реализации модели создано специальное программное обеспечение.



- Python 3.6 (numpy, pandas, sklearn, scipy, matplotlib, PyQt5);
- Windows & Linux

С использованием предложенной модели рассчитаны производители и качество 5 зашифрованных проб.

Результаты анализа зашифрованных проб

№	Ближайший кластер	Расстояние до центра (качество)	Среднее качество	Реальный производитель пробы
1	Ара NP(S) 12-40(10)	0.128	0.236	Ара NP(S) 12-40(10)
3	Ара NP(S) 12-40(10)	0.409	0.236	Ара NP(S) 12-40(10)
4	Met NPK(S) 0-20-20(5)	0.149	0.101	Met NPK(S) 0-20-20(5)
5	Met NPK(S) 0-20-20(5)	0.083	0.101	Met NPK(S) 0-20-20(5)

Дополнительно автоматически рассчитываются среднее расстояния до центра кластера - разбросе свойств внутри определенного типа удобрения (качество).

Выводы

1. Описаны основные свойства промышленно производимых удобрений в точки зрения «качества».
2. Исследована возможность многомерного моделирования производителя и «качества» минеральных удобрений.
3. Предложена модель кластеризации минеральных удобрений на основе ЭД РФА метода и проекции на метода проекции на главные с k-means.
4. В качестве метрик качества использована линейная комбинация Евклидова расстояния до центра кластера и критерий χ^2 .
5. Проведен эксперимент, показавший правильность выбранного подхода.