

## **ИССЛЕДОВАНИЕ ПРИМЕНИМОСТИ КОРРЕЛЯЦИОННОГО АНАЛИЗА ДЛЯ АВТОМАТИЗАЦИИ АНАЛИЗА ЭНЕРГЕТИЧЕСКИХ СПЕКТРОВ**

**Степанова А.В.**

**Научный руководитель – Сиротин Э.Е.**

***Сибирский федеральный университет***

Измерения параметров качества руды с использованием оборудования «Радос» выявили влияние расстояния от датчика до исследуемого объекта на результаты измерения параметров качества руды. В результате, не смотря на то, что метод измерения хорошо себя зарекомендовал на рудах мелкого класса (где расстояние между датчиком и объектом практически неизменно), он не может быть применим для руд высокого класса, ввиду больших погрешностей при замерах.

Следовательно, требуется новая обработка исходных данных, где расстояние от датчика до объекта – один из факторов, влияющих на качество получения исходных данных.

Уйти от высотной зависимости полностью нельзя, но можно уменьшить ее влияние. С этой целью и проводится данное исследование.

. Руда подвергается рентгеновскому облучению. В результате происходит возбуждение атомов, входящих в состав руды. Вторичное излучение улавливается детектором и измеряется энергия квантов, входящих в состав повторного излучения. Спектр представляет собой гистограмму распределения квантов от их энергий, зарегистрированных за единицу времени.

Энергетический спектр может включать в себя до 2048 каналов (зависит от характеристик устройств регистрации спектра), каждый из которых соответствует определенному диапазону энергии квантов, а значение в каждом канале соответствует количеству зарегистрированных квантов по этому диапазону за определенный промежуток времени. Путем анализа энергетического спектра можно определить величину концентрации определенных химических элементов в исследуемом образце/

Ниже представлена диаграмма спектров с проб руды с содержанием цинка 12,6% (рисунок 1), взятые с рудника в г. Зырянковске. Расстояние от измерительного прибора до объекта измерения на диаграмме 40, 34, 28 и 22 см. Из диаграммы видно, что значения спектров в зависимости от высоты сделанных замеров (расстояния от измерительного прибора до объекта измерения) различаются.

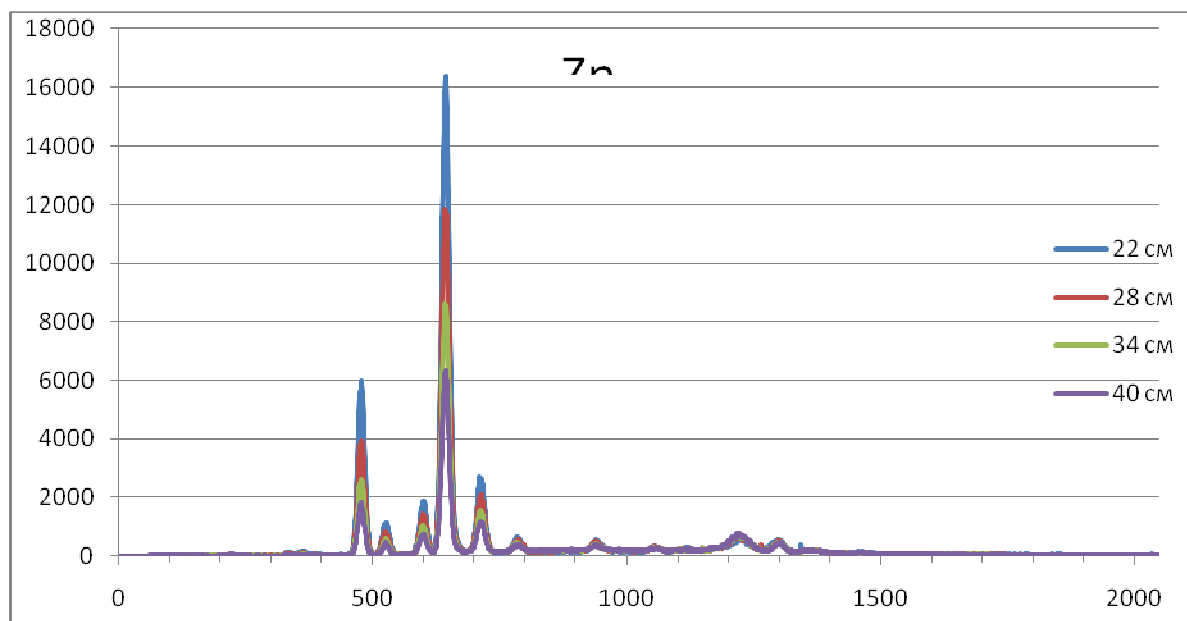


Рисунок 1. Диаграмма спектров проб с рудника с содержанием цинка 12,6%.

Следовательно, можно сделать вывод о том что высотная зависимость существует и требуется проводить дальнейшие исследования, чтобы узнать характер этой зависимости с целью корректировки данных, получаемых при замерах.

#### Выбор коэффициента корреляции.

В виду того, что нам необходимо проводить корреляционный анализ, используя континуальные шкалы, необходимо подробнее остановиться на коэффициенте корреляции Браве-Пирсона.

Коэффициент корреляции Браве-Пирсона ( $r_{xy}^p$ ) относится к параметрическим коэффициентам и для практических расчетов вычисляется по формуле:

$$r_{xy}^p = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}}$$

Из формулы видно, что для вычисления  $r_{xy}^p$  необходимо найти средние значения признаков X и Y, а также отклонения каждого статистического данного от его среднего

$(x_i - \bar{x}), (y_i - \bar{y})$ . Зная эти значения, находятся суммы  $\sum(x_i - \bar{x})(y_i - \bar{y})$ ,  $\sum(x_i - \bar{x})^2$ ,  $\sum(y_i - \bar{y})^2$ .

Затем, вычислив значение  $r_{xy}^p$ , необходимо определить достоверность найденного коэффициента корреляции, сравнив его фактическое значение с табличным для  $f = n - 2$ . Если  $r_{ф} \geq r_{ст}$ , то можно говорить о том, что между признаками наблюдается достоверная взаимосвязь. Если  $r_{ф} < r_{ст}$ , то между признаками наблюдается недостоверная корреляционная взаимосвязь.

При обзоре коэффициентов корреляции, опираясь на характер информации, которую нам предстоит обрабатывать, для проведения дальнейшего исследования был выбран коэффициент Браве-Пирсона.

## Вычисление коэффициентов корреляции.

Далее нужно вычислить непосредственно сами коэффициенты. Для этого необходима предварительная подготовка данных, необходимо выбрать инструмент для проведения исследования, который подойдет для вычисления большого объема данных.

### Исходные данные.

В качестве исходных данных были использованы замеры, сделанные на рудодобывающей станции в г. Зыряновске. Выбрано 26 проб. Каждая из проб сделана с четырех разных расстояний от датчика до объекта измерения (40 см., 34 см., 28 см., 22 см. соответственно). То есть, мы имеем 104 спектра в качестве исходных данных. Каждый спектр содержит 2048 каналов.

Исходные данные были сведены в таблицы, приведены в файлы формата excel. Эти таблицы очень большого размера, потому в данном отчете не имеется возможность их подробно представлять.

### Выбор инструмента исследования.

В качестве инструмента исследования нами был выбран matlab. Эта программа может вычислять большие объемы данных, в ней можно работать с матрицами, имеющими 2048 столбцов либо строк, в отличие, например, от excel, где максимум можно указывать 256 значений исходных данных для последующих расчетов.

Исходная таблица данных была разделена нами на 26 файлов формата excel, которые соответствуют числу проб, взятых на руднике г. Зыряновска. В каждом файле 2048 каналов энергетических спектров, полученных при замерах на четырех разных расстояниях от датчика до объекта измерения. Следовательно, размер исходной матрицы составил  $2048 \times 4$  (2048 строк  $\times$  4 столбца).

Далее нами был создан вектор  $v$ , который содержит в себе значения высот:

```
>> v = [40, 34, 28, 22]
v =
    40    34    28    22
```

Исходную матрицу было необходимо транспонировать, для того, чтобы соблюсти правила матричных операций при умножении ее элементов:

```
>> datac = data'
```

В matlab существует множество стандартных функций, которые можно использовать для расчетов различного характера. Возможно также написать собственную функцию, так называемую программу. Это необходимо в случае, если есть множество последовательных операций, либо нужно реализовать циклы, либо, если присутствуют операции с условиями.

Не смотря на то, что в matlab есть стандартная функция для подсчета парных коэффициентов корреляции:

$R = \text{corrcoef}(\text{data});$

Данная функция может вычислять корреляцию только непосредственно внутри матрицы относительно содержащихся в ней элементов. Поскольку эту функцию не представляется возможным использовать для дальнейшего исследования, мы столкнулись с тем, что эту задачу вычисления коэффициентов корреляции придется решать иным образом. С этой целью на языке matlab нами была написана своя функция для вычисления коэффициентов корреляции:

Для нашего исследования необходимо провести корреляцию между матрицей со значениями каналов и вектором, содержащим значения высот. Для этого нужно, чтобы каждый из 2048 каналов, который имеет 4 значения на разных высотах прокоррелировал с вектором, содержащим значения расстояния от датчика до объекта измерения. С этой целью, используя программирование в matlab, мы разбили каждую матрицу с исходными данными на 2048 векторов, содержащих в себе 4 элемента значений. В последствии был вычислен коэффициент корреляции между вектором из матрицы с каналами и вектором со значениями высот. Для одной пробы 2048 таких вычислений, что соответствует числу каналов в спектре. Следовательно, получается 2048 коэффициентов парной корреляции для каждой из проб (рисунок 5). Число коэффициентов соответствует числу каналов в каждом спектре. Таким же образом действовали на остальных 25-ти пробах.

#### Анализ полученных данных

С целью наглядного представления корреляционной зависимости между значениями каналов в энергетических спектрах и значениями расстояния от датчика до объекта измерения, был построен график значений коэффициентов корреляции на спектре (рисунок 2).

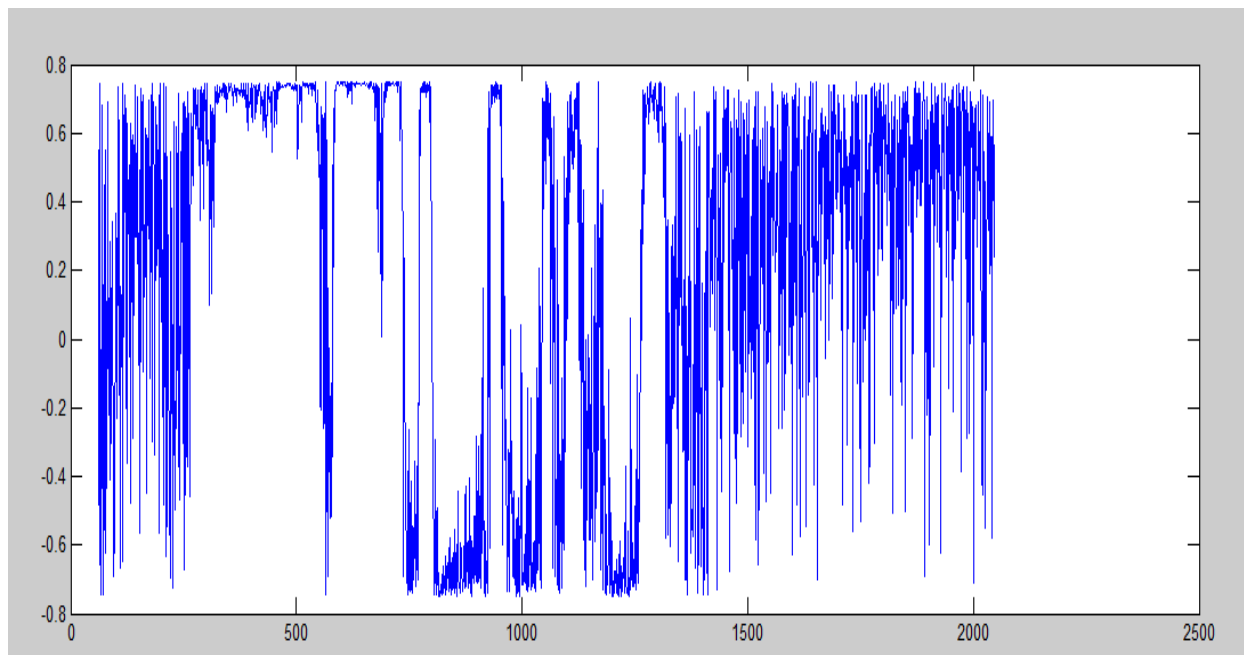


Рисунок 2. График корреляционной зависимости элементов 12 пробы.

По графику можно заметить, что корреляционная зависимость существует, но она различна для разных каналов энергетических спектров. Можно также увидеть, что

между отдельными каналами существует обратная корреляционная зависимость. Возможно, что при дальнейшем исследовании нам удастся выявить некую закономерность относительно полученных данных с обратной зависимостью.

В дальнейшем предстоит выяснить на каких каналах существует значительная корреляционная зависимость, и как значения в этих каналах влияют на исходные данные, получаемые при исследовании содержания полезных веществ во время просвечивания руды при помощи рентгеноспектрального анализа.

Полученный нами промежуточный результат в дальнейшем будет рассмотрен и проинтерпретирован совместно со специалистами ООО «РАДОС». Предстоит разработка методики – снижения влияния высотной зависимости при рентгеноспектральном анализе.

#### Вывод.

На данном этапе работы нами были подготовлены исходные данные для дальнейшего вычисления коэффициентов корреляции с использованием matlab, получена матрица коэффициентов корреляции.

Во время интерпретации данных на данном этапе исследования можно сказать, что явно видна зависимость между значениями отдельными каналами и значением расстояния от датчика до объекта измерения. Но также можно отметить, что эта зависимость проявляется по-разному между отдельными каналами спектров. Этот факт может позволить снизить влияние высотной зависимости.